

08/30/00

FRANK J. BOGANO  
DERICK E. ALLEN  
ROBERT M. BAIN  
JAMES J. BARTA, JR.  
G. HARLEY BLOSSER  
JOHN M. BODENHAUSEN  
RICHARD L. BRIDGE  
JAMES E. BUTLER, PH.D  
SARAH J. CHICKOS  
J. BENNETT CLARK  
JENNIFER E. COOK  
DAVID E. CRAWFORD, JR.  
MATTHEW L. CUTLER  
JAMES E. DAVIS  
KATHRYN J. DOTY  
ROBERT M. EVANS, JR.  
PAUL I.J. FLEISCHUT  
MICHAEL E. GODAR  
CHRISTOPHER M. GOFF  
DAVID M. GRYTE  
DAVID W. HARLAN

EDWARD J. HEJLEK  
KAREN Y. HUI  
KURT F. JAMES  
VINCENT M. KEIL  
ANTHONY R. KINNEY  
BRIAN P. KLEIN  
WILLIAM E. LAHEY  
PAUL A. MADDOCK  
MICHAEL G. MUNSELL  
DEBRA D. NYE  
KATHLEEN M. PETRILLO  
LAURA R. POLCYN  
KEITH A. RABENBERG  
STEVEN M. RITCHEY  
JOHN K. ROEDEL, JR.  
JOSEPH A. SCHAPER  
RICHARD A. SCHUTH  
MEG MARSHALL THOMAS  
MICHAEL J. THOMAS  
DONALD W. TUEGEL  
SCOTT A. WILLIAMS

SENNIGER, POWERS, LEAVITT & ROEDEL  
ATTORNEYS AT LAW

ONE METROPOLITAN SQUARE  
16TH FLOOR  
ST. LOUIS, MISSOURI 63102  
314-231-5400

August 30, 2000

FACSIMILE 314-231-4342  
<http://www.senniger.com>

PATENTS, TRADEMARKS, COPYRIGHTS  
AND RELATED MATTERS

OF COUNSEL  
IRVING POWERS  
DONALD G. LEAVITT  
RICHARD G. HEYWOOD  
ROBERT J. LEWIS

STUART N. SENNIGER  
(1921-1997)

1675 U.S. PTO  
09/651651  
08/30/00

UTILITY PATENT APPLICATION TRANSMITTAL  
(new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket Number: MTC 6718  
First Named Inventor: Michael Lassner  
Express Mail Label Number: EL615274183US

TO: Assistant Commissioner for Patents  
Box Patent Application  
Washington, D.C. 20231

APPLICATION ELEMENTS

1. ☒ Fee Transmittal Form  
(original and duplicate)
2. ☒ Pages of Application [Total Pages 73]
3. ☒ Drawings [Total Sheets 7]
4. Oath or Declaration [Total Pages 4]
  - a. ☐ Newly executed (original or copy)  
☒ New (unexecuted)
  - b. ☐ Copy from a prior application  
(for continuation/divisional with  
Box 17 completed)
    - i. ☐ DELETION OF INVENTOR(s)  
Signed statement attached  
deleting inventor(s) named  
in prior application.
5. ☐ Incorporation By Reference  
(useable if Box 4b is marked)  
The entire disclosure of the prior application, from which a  
copy of the oath or declaration is supplied under Box 4b, is  
considered as being part of the disclosure of the  
accompanying application and is hereby incorporated by  
reference therein.

6. ☐ Microfiche Computer Program (Appendix)
7. ☒ Nucleotide and/or Amino Acid Sequence Submission  
(if applicable, all necessary)
- a. ☒ Computer Readable Copy
  - b. ☒ Paper Copy (identical to computer copy)
  - c. ☒ Statement verifying identity of above copies

**ACCOMPANYING APPLICATION PARTS**

8. ☐ Assignment Papers (cover sheet & document(s))
9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney
10. ☐ English Translation Document (if applicable)
11. ☐ IDS with PTO-1449 ☐ Copies of IDS Citations
12. ☐ Preliminary Amendment
13. ☒ Return Receipt Postcard
14. ☐ Small Entity Statement(s)  
☐ Statement filed in prior application; status still proper and desired
15. ☐ Certified Copy of Priority Document(s) if foreign priority is claimed
16. ☐ Other: \_\_\_\_\_

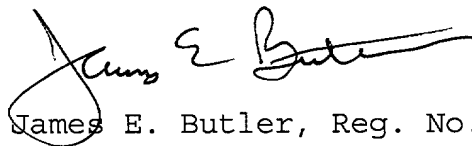
**IF A CONTINUING APPLICATION, CHECK APPROPRIATE  
BOXES AND SUPPLY THE REQUISITE INFORMATION**

17. ☐ Continuation ☐ Divisional ☐ Continuation-in-Part  
of prior application No.: \_\_\_\_\_
- ☒ Complete Application based on provisional Application  
No. 60/152,493 filed August 30, 1999

**CORRESPONDENCE ADDRESS**

18. Correspondence Address: Customer Number 321  
Attention: James E. Butler, Ph.D.

Respectfully submitted,

  
James E. Butler, Reg. No. 40,931

JEBmkd

Table 1. Demographic characteristics of the study population	
Age (years)	65.5 ± 1.2
Gender (male/female)	10/10
Education (years)	12.5 ± 0.5
Occupation (white/blue)	10/10
Marital status (married/divorced/widowed)	10/10/0
Smoking status (smoker/nonsmoker)	10/10
Alcohol consumption (yes/no)	10/10
Comorbidities (hypertension/diabetes/cholesterol)	10/10/10
Medication (antihypertensive/antidiabetic/anticholesterol)	10/10/10
Family history (hypertension/diabetes/cholesterol)	10/10/10
Physical activity (yes/no)	10/10
Stress level (low/moderate/high)	10/10/10
Social support (yes/no)	10/10
Health status (good/fair/poor)	10/10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction (yes/no)	10/10
Overall health (yes/no)	10/10
Physical health (yes/no)	10/10
Mental health (yes/no)	10/10
Social health (yes/no)	10/10
Emotional health (yes/no)	10/10
Financial health (yes/no)	10/10
Environmental health (yes/no)	10/10
Overall well-being (yes/no)	10/10
Life expectancy (yes/no)	10/10
Quality of life (yes/no)	10/10
Life satisfaction	

Application Number Not Yet Assigned  
Filed Herewith  
First Named Inventor Michael Lassner  
Attorney Docket Number MTC 6718

## METHOD OF PAYMENT

1.     [ ]     The Commissioner is hereby authorized to charge the indicated fees to Deposit Account No. 19-1345, in the name of Senniger, Powers, Leavitt & Roedel.  
  
          [ ]     The Commissioner is hereby authorized to charge any additional filing and claim fees under 37 CFR 1.16 and application processing fees under 37 CFR 1.17 to Deposit Account No. 19-1345, in the name of Senniger, Powers, Leavitt & Roedel.
2.     [X]     Check Enclosed. The Commissioner is hereby authorized to charge any under payment or credit any over payment to Deposit Account No. 19-1345, in the name of Senniger, Powers, Leavitt & Roedel.

### FEE CALCULATION

1. [X] BASIC FILING FEE \$ 690.00 (Type: Formal)  
Entity Status: Large
2. [X] CLAIM FEE \$ 4,490.00  
Total Claims 138  
Independent Claims 30  
Multiple Dependent Claims 18
3. [ ] ADDITIONAL FEES \$ \_\_\_\_\_  
[ ] Surcharge - late filing fee or oath  
[ ] Surcharge - late provisional filing fee or cover sheet  
[ ] Extension for reply within \_\_\_\_\_ month  
[ ] Notice of Appeal  
[ ] Filing a Brief in Support of an appeal  
[ ] Request for Reexamination  
[ ] Petitions to the Commissioner  
[ ] Submission of Information Disclosure Statement  
[ ] Recording each patent assignment per property  
[ ] Other: \_\_\_\_\_

**TOTAL AMOUNT OF PAYMENT**      \$ 5,180.00

James E. Butler, Ph.D., Reg. No. 40,931 8/30/00  
Date

**PLANT STEROL ACYLTRANSFERASES****CROSS-REFERENCE TO RELATED APPLICATIONS**

- 5           This application claims priority to U.S. provisional application Serial No. 60/152,493, filed August 30, 1999 and herein incorporated by reference in its entirety for all purposes.

**BACKGROUND**

10   Technical Field

The present invention is directed to plant acyltransferase-like nucleic acid and amino acid sequences and constructs, and methods related to their use in altering sterol composition and/or content, and oil composition and/or content in host cells and plants.

15   Related Art

Through the development of plant genetic engineering techniques, it is now possible to produce transgenic varieties of plant species to provide plants which have novel and desirable characteristics. For example, it is now possible to genetically engineer plants for tolerance to environmental stresses, such as resistance to pathogens and tolerance to herbicides. It is also possible to improve the nutritional characteristics of the plant, for example to provide improved fatty acid, carotenoid, sterol and tocopherol compositions. However, the number of useful nucleotide sequences for the engineering of such characteristics is thus far limited.

There is a need for improved means to obtain or manipulate compositions of sterols from biosynthetic or natural plant sources. The ability to increase sterol production or alter the sterol compositions in plants may provide for novel sources of sterols for use in human and animal nutrition.

Sterol biosynthesis branches from the farnesyl diphosphate intermediate in the isoprenoid pathway. Sterol biosynthesis occurs via a mevalonate dependent pathway in mammals and higher plants (Goodwin, (1981) *Biosynthesis of Isoprenoid Compounds*, vol 1 (Porter, J.W. & Spurgeon, S.L., eds) pp.443-480, John Wiley and Sons, New York), while in green algae sterol biosynthesis is thought to occur via a mevalonate independent pathway (Schwender, *et al.* (1997) *Physiology, Biochemistry, and Molecular Biology of Plant Lipids*, (Williams, J.P., Khan, M.U., and Lem, N.W., eds) pp. 180-182, Kluwer Academic Publishers, Norwell, MA).

The solubility characteristics of sterol esters suggests that this is the storage form of sterols (Chang, *et al.*, (1997) *Annu. Rev. Biochem.*, 66:613-638). Sterol O-acyltransferase enzymes such as acyl CoA:cholesterol acyltransferase (ACAT) and lecithin:cholesterol acyltransferase (LCAT) catalyze the formation of cholesterol esters, and thus are key to controlling the intracellular cholesterol storage. In yeast, it has been reported that overexpression of *LRO1*, a homolog of human LCAT, and phospholipid:diacylglycerol acyltransferase increased lipid synthesis (Oelkers *et al.*, (2000) *J. Biol. Chem.*, 26:15609-15612; Dahlqvist *et al.*, (2000) *Proc. Natl. Acad. Sci. USA*, 97:6487-6492).

The characterization of various acyltransferase proteins is useful for the further study of plant sterol synthesis systems and for the development of novel and/or alternative sterol sources. Studies of plant mechanisms may provide means to further enhance, control, modify, or otherwise alter the sterol composition of plant cells. Furthermore, such alterations in sterol content and/or composition may provide a means for obtaining tolerance to stress and insect damage. Of particular interest are the nucleic acid sequences of genes encoding proteins which may be useful for applications in genetic engineering.

### SUMMARY OF THE INVENTION

The present invention is directed to lecithin:cholesterol acyltransferase-like polypeptides (also referred to herein as LCAT) and acyl CoA:cholesterol acyltransferase-like polypeptides (also referred to herein as ACAT). In particular the invention is related to polynucleotides encoding such sterol:acyltransferases, polypeptides encoded by such polynucleotides, and the use of such polynucleotides to alter sterol composition and oil production. The polynucleotides of the present invention include those derived from plant sources.

One aspect of the invention, therefore, is an isolated nucleic acid sequence encoding a plant lecithin:cholesterol acyltransferase-like polypeptide, a fragment of a plant lecithin:cholesterol acyltransferase-like polypeptide, a plant acyl CoA:cholesterol acyltransferase-like polypeptide or a fragment of a plant acyl CoA:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence consisting essentially of SEQ ID NO: 2, 4, 6, 8, 10-29, 43-51, 73 or 75. Also provided is an isolated nucleic acid sequence consisting of SEQ ID NO: 2, 4, 6, 8, 10-29, 43-51, 73 or 75.

Still another aspect provides an isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 3 or SEQ ID NO: 3 with at least one conservative amino acid substitution; SEQ ID NO: 2; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 2; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 2; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 2 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Still another aspect provides an isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3' where X is a hydrogen, Y is a hydrogen or a metal, R<sub>1</sub> and R<sub>2</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 3 or SEQ ID NO: 3 with at least one conservative amino acid substitution; SEQ ID NO: 2; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 2; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 2; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 2 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 5 or SEQ ID NO: 5 with at least one conservative amino acid substitution; SEQ ID NO: 4; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 4; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 4; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 4 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3' where X is a hydrogen, Y is a hydrogen or a metal, R<sub>1</sub> and R<sub>2</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is selected from the group consisting of an isolated polynucleotide encoding a polypeptide of

SEQ ID NO: 5 or SEQ ID NO: 5 with at least one conservative amino acid substitution; SEQ ID NO: 4; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 4; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 4; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 4 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO:7 or SEQ ID NO: 7 with at least one conservative amino acid substitution; SEQ ID NO: 6; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 6; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 6; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 6 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3' where X is a hydrogen, Y is a hydrogen or a metal, R<sub>1</sub> and R<sub>2</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>3</sub> is selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 7 or SEQ ID NO: 7 with at least one conservative amino acid substitution; SEQ ID NO: 6; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 6; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 6; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 6 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO:9 or SEQ ID NO: 9 with at least one conservative amino acid substitution; SEQ ID NO: 8; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 8; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 8; an isolated polynucleotide

complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 8 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3' where X is a hydrogen, Y is a hydrogen or a metal, R<sub>1</sub> and R<sub>2</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 9 or SEQ ID NO: 9 with at least one conservative amino acid substitution; SEQ ID NO: 8; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 8; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 8; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 8 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 74 or SEQ ID NO: 74 with at least one conservative amino acid substitution; SEQ ID NO: 73; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 73; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 73; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 73 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3' where X is a hydrogen, Y is a hydrogen or a metal, R<sub>1</sub> and R<sub>2</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 74 or SEQ ID NO: 74 with at least one conservative amino acid substitution; SEQ ID NO: 73; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 73; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 73; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under



stringent conditions to SEQ ID NO: 73 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO:76 or SEQ ID NO: 76 with at least one conservative amino acid substitution; SEQ ID NO: 75; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 75; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 75; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 75 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3' where X is a hydrogen, Y is a hydrogen or a metal, R<sub>1</sub> and R<sub>2</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is selected from the group consisting of an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 76 or SEQ ID NO: 76 with at least one conservative amino acid substitution; SEQ ID NO: 75; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 75; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 75; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 75 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of SEQ ID NO: 42 or a degenerate variant thereof; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 42; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 42; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 42 and encodes an acyl CoA:cholesterol acyltransferase-like polypeptide.

Another aspect provides an isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3' where X is a hydrogen, Y is a hydrogen or a metal, R<sub>1</sub> and R<sub>2</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is

selected from the group consisting of SEQ ID NO: 42 or a degenerate variant thereof; an isolated polynucleotide that has at least 70%, 80%, 90%, or 95% sequence identity with SEQ ID NO: 42; an isolated polynucleotide of at least 10 amino acids that hybridizes under stringent conditions to SEQ ID NO: 42; an isolated polynucleotide complementary to any of the foregoing; and an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 42 and encodes a acyl CoA:cholesterol acyltransferase-like polypeptide.

Also provided is a recombinant nucleic acid construct comprising a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide and/or an acyl CoA:cholesterol acyltransferase-like polypeptide. In one embodiment, the sterol acyl transferases are plant sterol acyl transferases. In another embodiment, the recombinant nucleic acid constructs further comprises a termination sequence. The regulatory sequence can be a constitutive promoter, an inducible promoter, a developmentally regulated promoter, a tissue specific promoter, an organelle specific promoter, a seed specific promoter or a combination of any of the foregoing. Also provided is a plant containing this recombinant nucleic acid construct and the seed and progeny from such a plant. This recombinant nucleic acid construct can also be introduced into a suitable host cell to provide yet another aspect of the invention. If the host cell is a plant host cell, the cell can be used to generate a plant to provide another aspect of the invention. Further aspects include seed and progeny from such a plant.

Yet another aspect is a purified polypeptide comprising, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 74, SEQ ID NO: 76, or any of the preceding sequences with at least one conservative amino acid substitution.

Still another aspect provides a purified immunogenic polypeptide comprising at least 10 consecutive amino acids from an amino acid sequence selected from the group consisting of SEQ ID NO: 3, 5, 7, 9, 74, 76 and any of the preceding sequences containing at least one conservative amino acid substitution. Also provided are antibodies, either polyclonal or monoclonal, that specifically bind the preceding immunogenic polypeptides.

One aspect provides a method for producing a lecithin:cholesterol acyltransferase-like polypeptide or an acyl CoA:cholesterol acyltransferase-like polypeptide comprising culturing a host cell containing any recombinant nucleic acid construct of the present invention under condition permitting expression of said lecithin:cholesterol acyltransferase-like polypeptide or acyl CoA:cholesterol acyltransferase-like polypeptide.

Another aspect provides a method for modifying the sterol content of a host cell, comprising transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide and culturing said host cell under conditions wherein said host cell expresses a lecithin:cholesterol acyltransferase-like polypeptide such that said host cell has a modified sterol composition as compared to host cells without the recombinant construct.

An additional aspect is a method for modifying the sterol content of a host cell comprising transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide and culturing said host cell under conditions wherein said host cell expresses an acyl CoA:cholesterol acyltransferase-like polypeptide such that said host cell has a modified sterol composition as compared to host cells without the recombinant construct.

A further aspect is a plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in modified sterol composition of said plant as compared to the same plant without said recombinant construct.

Another aspect provides a plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in modified sterol composition of said plant as compared to the same plant without said recombinant construct.

In a further aspect is provided an oil obtained from any of the plants or host cells of the present invention.

In still another aspect is provided a method for producing an oil with a modified sterol composition comprising providing any of the plants or host cells of the present invention and extracting oil from the plant by any known method. Also provided is an oil produced by the preceding method.

Still another aspect provides a method for altering oil production by a host cell comprising, transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide and culturing the host cell under conditions wherein the host cell expresses a

lecithin:cholesterol acyltransferase-like polypeptide such that the host cell has an altered oil production as compared to host cells without the recombinant construct.

Another aspect provides a method for altering oil production by a host cell comprising, transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide and culturing the host cell under conditions wherein the host cell expresses an acyl CoA:cholesterol acyltransferase-like polypeptide such that the host cell has an altered oil production as compared to host cells without the recombinant construct.

Also provided is a plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in an altered production of oil by said plant as compared to the same plant without said recombinant construct.

In a further aspect is provided a plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in an altered production of oil by said plant as compared to the same plant without said recombinant construct.

Additional aspects provide a food, food ingredient or food product comprising any oil, plant or host cell of the present invention; a nutritional or dietary supplement comprising any oil, plant or host cell of the present invention; and a pharmaceutical composition comprising any oil, plant or host cell of the present invention along with a suitable diluent, carrier or excipient.

Additional aspects will be apparent from the descriptions and examples that follow.

## BRIEF DESCRIPTION OF THE DRAWINGS

These and other features, aspects, and advantages of the present invention will become better understood with regard to the following description, appended claims and accompanying figures where:

Figure 1 shows an alignment of yeast, human and rat lecithin:cholesterol acyltransferase protein sequences with *Arabidopsis* LCAT1, LCAT2, LCAT3, and LCAT4 deduced amino acid sequences.

Figure 2 shows the results of NMR sterol ester analysis on T2 seed from plant expressing LCAT4 under the control of the napin promoter (pCGN9998).

Figure 3 shows the results of HPLC/MS sterol analysis on oil extracted from T2 seed from control lines (pCGN8640) and lines expressing LCAT3 (pCGN9968) under the control of the napin promoter.

Figure 4 shows the results of HPLC/MS sterol analysis on oil extracted from T2 seed from control lines (pCGN8640), and plant line expressing LCAT1 (pCGN9962), LCAT2 (pCGN9983), LCAT3 (pCGN9968), and LCAT4 (pCGN9998) under the control of the napin promoter. Additionally, data from 3 lines expressing LCAT4 under the control of the 35S promoter (pCGN9996) are shown.

Figure 5 shows the results of Nir analysis of the oil content of T2 seed from control lines (pCGN8640), and plant lines expressing LCAT1 (pCGN9962), LCAT2 (pCGN9983), and LCAT3 (pCGN9968) under the control of the napin promoter. Additionally, data from 16 lines expressing LCAT2 under the control of the 35S promoter (pCGN9981) are shown.

#### DETAILED DESCRIPTION

The following detailed description is provided to aid those skilled in the art in practicing the present invention. Even so, this detailed description should not be construed to unduly limit the present invention as modifications and variations in the embodiments discussed herein can be made by those of ordinary skill in the art without departing from the spirit or scope of the present inventive discovery.

All publications, patents, patent applications and other references cited in this application are herein incorporated by reference in their entirety as if each individual publication, patent, patent application or other reference were specifically and individually indicated to be incorporated by reference.

The present invention relates to lecithin:cholesterol acyltransferase, particularly the isolated nucleic acid sequences encoding lecithin:cholesterol-like polypeptides (LCAT) from plant sources and acyl CoA:cholesterol:acyltransferase, particularly the isolated nucleic acid sequences encoding acyl CoA:cholesterol acyltransferase-like polypeptides (ACAT) from plant sources. Lecithin:cholesterol acyltransferase-like as used herein includes any nucleic acid sequence encoding an amino acid sequence from a plant source, such as a protein, polypeptide or peptide, obtainable from a cell source, which demonstrates the ability to utilize

lecithin (phosphatidyl choline) as an acyl donor for acylation of sterols or glycerides to form esters under enzyme reactive conditions along with such proteins polypeptides and peptides. Acyl CoA:cholesterol acyltransferase-like as used herein includes any nucleic acid sequence encoding an amino acid sequence from a plant source, such as a protein, polypeptide or peptide, obtainable from a cell source, which demonstrates the ability to utilize acyl CoA as an acyl donor for acylation of sterols or glycerides to form esters under enzyme reactive conditions along with such proteins polypeptides and peptides. By "enzyme reactive conditions" is meant that any necessary conditions are available in an environment (i.e., such factors as temperature, pH, lack of inhibiting substances) which will permit the enzyme to function.

The term "sterol" as applied to plants refers to any chiral tetracyclic isopentenoid which may be formed by cyclization of squalene oxide through the transition state possessing stereochemistry similar to the *trans-syn-trans-anti-trans-anti* configuration, for example, protosteroid cation, and which retains a polar group at C-3 (hydroxyl or keto), an *all-trans-anti* stereochemistry in the ring system, and a side-chain 20R-configuration (Parker, *et al.* (1992) In Nes, *et al.*, Eds., *Regulation of Isopentenoid Metabolism*, ACS Symposium Series No. 497, p. 110; American Chemical Society, Washington, D.C.).

Sterols may or may not contain a C-5-C-6 double bond, as this is a feature introduced late in the biosynthetic pathway. Sterols contain a C<sub>8</sub>-C<sub>10</sub> side chain at the C-17 position.

The term "phytosterol," which applies to sterols found uniquely in plants, refers to a sterol containing a C-5, and in some cases a C-22, double bond. Phytosterols are further characterized by alkylation of the C-17 side-chain with a methyl or ethyl substituent at the C-24 position. Major phytosterols include, but are not limited to, sitosterol, stigmasterol, campesterol, brassicasterol, etc. Cholesterol, which lacks a C-24 methyl or ethyl side-chain, is found in plants, but is not unique thereto, and is not a "phytosterol."

"Phytostanols" are saturated forms of phytosterols wherein the C-5 and, when present, C-22 double bond(s) is (are) reduced, and include, but are not limited to, sitostanol, campestanol, and 22-dihydrobrassicastanol.

"Sterol esters" are further characterized by the presence of a fatty acid or phenolic acid moiety rather than a hydroxyl group at the C-3 position.

The term "sterol" includes sterols, phytosterols, phytosterol esters, phytostanols, and phytostanol esters.

The term "sterol compounds" includes sterols, phytosterols, phytosterol esters, phytostanols, and phytostanol esters.

The term "phytosterol compound" refers to at least one phytosterol, at least one phytosterol ester, or a mixture thereof.

5 The term "phytostanol compound" refers to at least one phytostanol, at least one phytostanol ester, or a mixture thereof.

The term "glyceride" refers to a fatty acid ester of glycerol and includes mono-, di-, and tri- acylglycerols.

10 As used herein, "recombinant construct" is defined either by its method of production or its structure. In reference to its method of production, e.g., a product made by a process, the process is use of recombinant nucleic acid techniques, e.g., involving human intervention in the nucleotide sequence, typically selection or production. Alternatively, in terms of structure, it can be a sequence comprising fusion of two or more nucleic acid sequences which are not naturally contiguous or operatively linked to each other

15 As used herein, "regulatory sequence" means a sequence of DNA concerned with controlling expression of a gene; e.g. promoters, operators and attenuators. A "heterologous regulatory sequence" is one which differs from the regulatory sequence naturally associated with a gene.

20 As used herein, "polynucleotide" and "oligonucleotide" are used interchangeably and mean a polymer of at least two nucleotides joined together by a phosphodiester bond and may consist of either ribonucleotides or deoxynucleotides.

As used herein, "sequence" means the linear order in which monomers appear in a polymer, for example, the order of amino acids in a polypeptide or the order of nucleotides in a polynucleotide.

25 As used herein, "polypeptide", "peptide", and "protein" are used interchangeably and mean a compound that consist of two or more amino acids that are linked by means of peptide bonds.

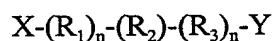
30 As used herein, the terms "complementary" or "complementarity" refer to the pairing of bases, purines and pyrimidines, that associate through hydrogen bonding in double stranded nucleic acids. For example, the following base pairs are complementary: guanine and cytosine; adenine and thymine; and adenine and uracil. The terms, as used herein, include complete and partial complementarity.

**Isolated proteins, Polypeptides and Polynucleotides**

A first aspect of the present invention relates to isolated LCAT polynucleotides. The polynucleotide sequences of the present invention include isolated polynucleotides that encode the polypeptides of the invention having a deduced amino acid sequence selected from the group of sequences set forth in the Sequence Listing and to other polynucleotide sequences closely related to such sequences and variants thereof.

The invention provides a polynucleotide sequence identical over its entire length to each coding sequence as set forth in the Sequence Listing. The invention also provides the coding sequence for the mature polypeptide or a fragment thereof, as well as the coding sequence for the mature polypeptide or a fragment thereof in a reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, pro-, or prepro- protein sequence. The polynucleotide can also include non-coding sequences, including for example, but not limited to, non-coding 5' and 3' sequences, such as the transcribed, untranslated sequences, termination signals, ribosome binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and additional coding sequence that encodes additional amino acids. For example, a marker sequence can be included to facilitate the purification of the fused polypeptide. Polynucleotides of the present invention also include polynucleotides comprising a structural gene and the naturally associated sequences that control gene expression.

The invention also includes polynucleotides of the formula:



wherein, at the 5' end, X is hydrogen, and at the 3' end, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any nucleic acid residue,  $n$  is an integer between 0 and 3000, preferably between 1 and 1000 and  $R_2$  is a nucleic acid sequence of the invention, particularly a nucleic acid sequence selected from the group set forth in the Sequence Listing and preferably SEQ ID NOs: 2, 4, 6, 8, 10-29, 33, 42-51, 73 and 75. In the formula,  $R_2$  is oriented so that its 5' end residue is at the left, bound to  $R_1$ , and its 3' end residue is at the right, bound to  $R_3$ . Any stretch of nucleic acid residues denoted by either R group, where  $R$  is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer.

The invention also relates to variants of the polynucleotides described herein that encode for variants of the polypeptides of the invention. Variants that are fragments of the



polynucleotides of the invention can be used to synthesize full-length polynucleotides of the invention. Preferred embodiments are polynucleotides encoding polypeptide variants wherein 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues of a polypeptide sequence of the invention are substituted, added or deleted, in any combination. Particularly preferred are substitutions, additions, and deletions that are silent such that they do not alter the properties or activities of the polynucleotide or polypeptide.

Further preferred embodiments of the invention that are at least 50%, 60%, or 70% identical over their entire length to a polynucleotide encoding a polypeptide of the invention, and polynucleotides that are complementary to such polynucleotides. More preferable are polynucleotides that comprise a region that is at least 80% identical over its entire length to a polynucleotide encoding a polypeptide of the invention and polynucleotides that are complementary thereto. In this regard, polynucleotides at least 90% identical over their entire length are particularly preferred, those at least 95% identical are especially preferred. Further, those with at least 97% identity are highly preferred and those with at least 98% and 99% identity are particularly highly preferred, with those at least 99% being the most highly preferred.

Preferred embodiments are polynucleotides that encode polypeptides that retain substantially the same biological function or activity as determined by the methods described herein as the mature polypeptides encoded by the polynucleotides set forth in the Sequence Listing.

The invention further relates to polynucleotides that hybridize to the above-described sequences. In particular, the invention relates to polynucleotides that hybridize under stringent conditions to the above-described polynucleotides. An example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising 50% formamide, 5x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/milliliter denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at approximately 65°C. Also included are polynucleotides that hybridize under a wash stringency of 0.1X SSC or 0.1X SSPE (at 50°C. Other hybridization and wash conditions are well known and are exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, NY (1989), particularly Chapter 11.

The invention also provides a polynucleotide consisting essentially of a polynucleotide sequence obtainable by screening an appropriate library containing the

complete gene for a polynucleotide sequence set forth in the Sequence Listing under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence or a fragment thereof; and isolating said polynucleotide sequence. Methods for screening libraries are well known in the art and can be found for example in Sambrook, *et al.*,

5 *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, NY (1989), particularly Chapter 8 and Ausubel *et al.*, *Short Protocols in Molecular Biology*, 3<sup>rd</sup> ed, Wiley and Sons, 1995, chapter 6. Nucleic acid sequences useful for obtaining such a polynucleotide include, for example, probes and primers as described herein and in particular SEQ ID NO: 2, 4, 6, 8, 10-29, 33, 42-51, 73 and 75. These sequences are particularly useful  
10 in screening libraries obtained from *Arabidopsis*, soybean and corn for sequences encoding lecithin:cholesterol acyltransferase and lecithin:cholesterol acyltransferase-like polypeptides and for screening libraries for sequences encoding acyl CoA:cholesterol acyl transferase and acyl CoA:cholesterol acyl transferase-like polypeptides.

As discussed herein regarding polynucleotide assays of the invention, for example,  
15 polynucleotides of the invention can be used as a hybridization probe for RNA, cDNA, or genomic DNA to isolate full length cDNAs or genomic clones encoding a polypeptide and to isolate cDNA or genomic clones of other genes that have a high sequence similarity to a polynucleotide set forth in the Sequence Listing and in particular SEQ ID NO: 2, 4, 6, 8, 10-29, 33, 42-51, 73 and 75. Such probes will generally comprise at least 15 bases. Preferably  
20 such probes will have at least 30 bases and can have at least 50 bases. Particularly preferred probes will have between 30 bases and 50 bases, inclusive.

The coding region of each gene that comprises or is comprised by a polynucleotide sequence set forth in the Sequence Listing may be isolated by screening using a DNA sequence provided in the Sequence Listing to synthesize an oligonucleotide probe. A labeled  
25 oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to identify members of the library which hybridize to the probe. For example, synthetic oligonucleotides are prepared which correspond to the LCAT EST sequences. The oligonucleotides are used as primers in polymerase chain reaction (PCR) techniques to obtain 5' and 3' terminal sequence of LCAT  
30 genes. Alternatively, where oligonucleotides of low degeneracy can be prepared from particular LCAT peptides, such probes may be used directly to screen gene libraries for LCAT gene sequences. In particular, screening of cDNA libraries in phage vectors is useful in such methods due to lower levels of background hybridization.

Typically, a LCAT sequence obtainable from the use of nucleic acid probes will show 60-70% sequence identity between the target LCAT sequence and the encoding sequence used as a probe. However, lengthy sequences with as little as 50-60% sequence identity may also be obtained. The nucleic acid probes may be a lengthy fragment of the nucleic acid sequence, or may also be a shorter, oligonucleotide probe. When longer nucleic acid fragments are employed as probes (greater than about 100 bp), one may screen at lower stringencies in order to obtain sequences from the target sample which have 20-50% deviation (i.e., 50-80% sequence homology) from the sequences used as probe.

Oligonucleotide probes can be considerably shorter than the entire nucleic acid sequence encoding an LCAT enzyme, but should be at least about 10, preferably at least about 15, and more preferably at least about 20 nucleotides. A higher degree of sequence identity is desired when shorter regions are used as opposed to longer regions. It may thus be desirable to identify regions of highly conserved amino acid sequence to design oligonucleotide probes for detecting and recovering other related LCAT genes. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be identified. (See, Gould, *et al.*, *PNAS USA* (1989) 86:1934-1938.).

Another aspect of the present invention relates to LCAT polypeptides. Such polypeptides include isolated polypeptides set forth in the Sequence Listing, as well as polypeptides and fragments thereof, particularly those polypeptides which exhibit LCAT activity and also those polypeptides which have at least 50%, 60% or 70% identity, preferably at least 80% identity, more preferably at least 90% identity, and most preferably at least 95% identity to a polypeptide sequence selected from the group of sequences set forth in the Sequence Listing, and also include portions of such polypeptides, wherein such portion of the polypeptide preferably includes at least 30 amino acids and more preferably includes at least 50 amino acids.

"Identity", as is well understood in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods including, but not limited to, those described in *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York (1988); *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin,

A.M. and Griffin, H.G., eds., Humana Press, New Jersey (1994); *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press (1987); *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., Stockton Press, New York (1991); and Carillo, H., and Lipman, D., *SIAM J Applied Math*, 48:1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available programs. Computer programs which can be used to determine identity between two sequences include, but are not limited to, GCG (Devereux, J., et al., *Nucleic Acids Research* 12(1):387 (1984); suite of five BLAST programs, three designed for nucleotide sequences queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology*, 12: 76-80 (1994); Birren, et al., *Genome Analysis*, 1: 543-559 (1997)). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH, Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.*, 215:403-410 (1990)). The well known Smith Waterman algorithm can also be used to determine identity.

Parameters for polypeptide sequence comparison typically include the following:

Algorithm: Needleman and Wunsch, *J. Mol. Biol.* 48:443-453 (1970)

Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl. Acad. Sci USA* 89:10915-10919 (1992)

Gap Penalty: 12

Gap Length Penalty: 4

A program which can be used with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison Wisconsin. The above parameters along with no penalty for end gap are the default parameters for peptide comparisons.

Parameters for polynucleotide sequence comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol. Biol.* 48:443-453 (1970)

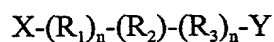
Comparison matrix: matches = +10; mismatches = 0

Gap Penalty: 50

Gap Length Penalty: 3

A program which can be used with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison Wisconsin. The above parameters are the default parameters for nucleic acid comparisons.

The invention also includes polypeptides of the formula:



wherein, at the amino terminus, X is hydrogen, and at the carboxyl terminus, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any amino acid residue,  $n$  is an integer between 0 and 1000, and  $R_2$  is an amino acid sequence of the invention, particularly an amino acid sequence selected from the group set forth in the Sequence Listing and preferably SEQ ID NOs: 3, 5, 7, 9, 74 and 76. In the formula,  $R_2$  is oriented so that its amino terminal residue is at the left, bound to  $R_1$ , and its carboxy terminal residue is at the right, bound to  $R_3$ . Any stretch of amino acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer.

Polypeptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising a sequence selected from the group of a sequence contained in SEQ ID NOs: 2, 4, 6, 8, 73 and 75.

The polypeptides of the present invention can be mature protein or can be part of a fusion protein.

Fragments and variants of the polypeptides are also considered to be a part of the invention. A fragment is a variant polypeptide which has an amino acid sequence that is entirely the same as part but not all of the amino acid sequence of the previously described polypeptides. The fragments can be "free-standing" or comprised within a larger polypeptide of which the fragment forms a part or a region, most preferably as a single continuous region. Preferred fragments are biologically active fragments which are those fragments that mediate activities of the polypeptides of the invention, including those with similar activity or improved activity or with a decreased activity. Also included are those polypeptides and polypeptide fragments that are antigenic or immunogenic in an animal, particularly a human and antibodies, either polyclonal or monoclonal that specifically bind the antigenic fragments.

In one preferred embodiment, such antigenic or immunogenic fragments comprise at least 10 consecutive amino acids from the amino acid sequences disclosed herein or such sequences with at least one conservative amino acid substitution. In additional embodiments, such antigenic or immunogenic fragments comprise at least 15, at least 25, at least 50 or at least 100 consecutive amino acids from the amino acid sequences disclosed herein or such sequences with at least one conservative amino acid substitution. Methods for the production of antibodies from polypeptides and polypeptides conjugated to carrier molecules are well known in the art and can be found for example in Ausubel et al., *Short Protocols in Molecular Biology*, 3<sup>rd</sup> ed., Wiley & Sons, 1995, particularly chapter 11.

5

15

20

25

Like amino acids can also be substituted on the basis of hydrophilicity. U.S. Patent No. 4,554,101 discloses that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological

property of the protein. The following hydrophilicity values have been assigned to amino acids: arginine/lysine (+3.0); aspartate/glutamate (+3.0  $\pm$ 1); serine (+0.3); asparagine/glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5  $\pm$ 1); alanine/histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine/isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). Thus, one amino acid in a peptide, polypeptide, or protein can be substituted by another amino acid having a similar hydrophilicity score and still produce a resultant protein having similar biological activity, i.e., still retaining correct biological function. In making such changes, amino acids having hydrophobic indices within  $\pm$ 2 are preferably substituted for one another, those within  $\pm$ 1 are more preferred, and those within  $\pm$ 0.5 are most preferred.

As outlined above, amino acid substitutions in the peptides of the present invention can be based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, etc. Exemplary substitutions that take various of the foregoing characteristics into consideration in order to produce conservative amino acid changes resulting in silent changes within the present peptides, etc., can be selected from other members of the class to which the naturally occurring amino acid belongs. Amino acids can be divided into the following four groups: (1) acidic amino acids; (2) basic amino acids; (3) neutral polar amino acids; and (4) neutral non-polar amino acids. Representative amino acids within these various groups include, but are not limited to: (1) acidic (negatively charged) amino acids such as aspartic acid and glutamic acid; (2) basic (positively charged) amino acids such as arginine, histidine, and lysine; (3) neutral polar amino acids such as glycine, serine, threonine, cysteine, cystine, tyrosine, asparagine, and glutamine; and (4) neutral non-polar amino acids such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine. It should be noted that changes which are not expected to be advantageous can also be useful if these result in the production of functional sequences.

Variants that are fragments of the polypeptides of the invention can be used to produce the corresponding full length polypeptide by peptide synthesis. Therefore, these variants can be used as intermediates for producing the full-length polypeptides of the invention.

The polynucleotides and polypeptides of the invention can be used, for example, in the transformation of host cells, such as plant cells, animal cells, yeast cells, bacteria, bacteriophage, and viruses, as further discussed herein.

The invention also provides polynucleotides that encode a polypeptide that is a mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids within the mature polypeptide (for example, when the mature form of the protein has more than one polypeptide chain). Such sequences can, for example, play a role in the processing of a protein from a precursor to a mature form, allow protein transport, shorten or lengthen protein half-life, or facilitate manipulation of the protein in assays or production. It is contemplated that cellular enzymes can be used to remove any additional amino acids from the mature protein.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. The inactive precursors generally are activated when the prosequences are removed. Some or all of the prosequences may be removed prior to activation. Such precursor protein are generally called proproteins.

#### **Preparation of Expression Constructs and Methods of Use**

Of interest is the use of the nucleotide sequences in recombinant DNA constructs to direct the transcription or transcription and translation (expression) of the acyltransferase sequences of the present invention in a host cell. Of particular interest is the use of the polynucleotide sequences of the present invention in recombinant DNA constructs to direct the transcription or transcription and translation (expression) of the acyltransferase sequences of the present invention in a host plant cell.

The expression constructs generally comprise a regulatory sequence functional in a host cell operably linked to a nucleic acid sequence encoding a lecithin:cholesterol acyltransferase-like polypeptide or acyl CoA:cholesterol acyltransferase-like polypeptide of the present invention and a transcriptional termination region functional in a host plant cell. Of particular interest is the use of promoters (also referred to as transcriptional initiation regions) functional in plant host cells.

Those skilled in the art will recognize that there are a number of promoters which are functional in plant cells, and have been described in the literature including constitutive, inducible, tissue specific, organelle specific, developmentally regulated and environmentally regulated promoters. Chloroplast and plastid specific promoters, chloroplast or plastid functional promoters, and chloroplast or plastid operable promoters are also envisioned.



One set of promoters are constitutive promoters such as the CaMV35S or FMV35S promoters that yield high levels of expression in most plant organs. Enhanced or duplicated versions of the CaMV35S and FMV35S promoters are useful in the practice of this invention (Odell, *et al.* (1985) *Nature* 313:810-812; Rogers, U.S. Patent Number 5,378, 619). Other useful constitutive promoters include, but are not limited to, the mannopine synthase (*mas*) promoter, the nopaline synthase (*nos*) promoter, and the octopine synthase (*ocs*) promoter.

Useful inducible promoters include heat-shock promoters (Ou-Lee et al. (1986) *Proc. Natl. Acad. Sci. USA* 83: 6815; Ainley et al. (1990) *Plant Mol. Biol.* 14: 949), a nitrate-inducible promoter derived from the spinach nitrite reductase gene (Back et al. (1991) *Plant Mol. Biol.* 17: 9), hormone-inducible promoters (Yamaguchi-Shinozaki et al. (1990) *Plant Mol. Biol.* 15: 905; Kares et al. (1990) *Plant Mol. Biol.* 15: 905), and light-inducible promoters associated with the small subunit of RuBP carboxylase and LHCP gene families (Kuhlemeier et al. (1989) *Plant Cell* 1: 471; Feinbaum et al. (1991) *Mol. Gen. Genet.* 226: 449; Weisshaar et al. (1991) *EMBO J.* 10: 1777; Lam and Chua (1990) *Science* 248: 471; Castresana et al. (1988) *EMBO J.* 7: 1929; Schulze-Lefert et al. (1989) *EMBO J.* 8: 651).

In addition, it may also be preferred to bring about expression of the acyltransferase gene in specific tissues of the plant, such as leaf, stem, root, tuber, seed, fruit, etc., and the promoter chosen should have the desired tissue and developmental specificity. Examples of useful tissue-specific, developmentally-regulated promoters include fruit-specific promoters such as the E4 promoter (Cordes et al. (1989) *Plant Cell* 1:1025), the E8 promoter (Deikman et al. (1988) *EMBO J.* 7: 3315), the kiwifruit actinidin promoter (Lin et al. (1993) *PNAS* 90: 5939), the 2A11 promoter (Houck et al., U.S. Patent 4,943,674), and the tomato pZ130 promoter (U.S. Patents 5,175, 095 and 5,530,185); the  $\beta$ -conglycinin 7S promoter (Doyle et al. (1986) *J. Biol. Chem.* 261: 9228; Slighton and Beachy (1987) *Planta* 172: 356), and seed-specific promoters (Knutzon et al. (1992) *Proc. Natl. Acad. Sci. USA* 89: 2624; Bustos et al. (1991) *EMBO J.* 10: 1469; Lam and Chua (1991) *J. Biol. Chem.* 266: 17131; Stayton et al. (1991) *Aust. J. Plant. Physiol.* 18: 507). Fruit-specific gene regulation is discussed in U.S. Patent 5,753,475. Other useful seed-specific promoters include, but are not limited to, the napin, phaseolin, zein, soybean trypsin inhibitor, 7S, ADR12, ACP, stearyl-ACP desaturase, oleosin, *Lasquerella* hydroxylase, and barley aldose reductase promoters (Bartels (1995) *Plant J.* 7: 809-822), the EA9 promoter (U.S. Patent 5,420,034), and the Bce4 promoter (U.S. Patent 5,530,194). Useful embryo-specific promoters include the corn globulin 1 and oleosin promoters. Useful endosperm-specific promoters include the rice glutelin-1 promoter, the

promoters for the low-pI  $\beta$  amylase gene (Amy32b) (Rogers et al. (1984) *J. Biol. Chem.* 259: 12234), the high-pI  $\beta$  amylase gene (Amy 64) (Khurseed et al. (1988) *J. Biol. Chem.* 263: 18953), and the promoter for a barley thiol protease gene ("Aleurain") (Whittier et al. (1987) *Nucleic Acids Res.* 15: 2515).

5 Of particular interest is the expression of the nucleic acid sequences of the present invention from transcription initiation regions which are preferentially expressed in a plant seed tissue. Examples of such seed preferential transcription initiation sequences include those sequences derived from sequences encoding plant storage protein genes or from genes involved in fatty acid biosynthesis in oilseeds. Examples of such promoters include the 5'  
10 regulatory regions from such genes as napin (Kridl et al., *Seed Sci. Res.* 1:209:219 (1991)), phaseolin, zein, soybean trypsin inhibitor, ACP, stearyl-ACP desaturase, soybean  $\alpha'$  subunit of  $\beta$ -conglycinin (soy 7s, (Chen et al., *Proc. Natl. Acad. Sci.*, 83:8560-8564 (1986))) and oleosin. Seed-specific gene regulation is discussed in EP 0 255 378 B1 and U.S. Patents 5,420,034 and 5,608,152 . Promoter hybrids can also be constructed to enhance  
15 transcriptional activity (Hoffman, U.S. Patent No. 5,106,739), or to combine desired transcriptional activity and tissue specificity.

It may be advantageous to direct the localization of proteins conferring LCAT to a particular subcellular compartment, for example, to the mitochondrion, endoplasmic reticulum, vacuoles, chloroplast or other plastidic compartment. For example, where the  
20 genes of interest of the present invention will be targeted to plastids, such as chloroplasts, for expression, the constructs will also employ the use of sequences to direct the gene to the plastid. Such sequences are referred to herein as chloroplast transit peptides (CTP) or plastid transit peptides (PTP). In this manner, where the gene of interest is not directly inserted into the plastid, the expression construct will additionally contain a gene encoding a transit  
25 peptide to direct the gene of interest to the plastid. The chloroplast transit peptides may be derived from the gene of interest, or may be derived from a heterologous sequence having a CTP. Such transit peptides are known in the art. See, for example, Von Heijne et al. (1991) *Plant Mol. Biol. Rep.* 9:104-126; Clark et al. (1989) *J. Biol. Chem.* 264:17544-17550; della Cioppa et al. (1987) *Plant Physiol.* 84:965-968; Romer et al. (1993) *Biochem. Biophys. Res Commun.* 196:1414-1421; and, Shah et al. (1986) *Science* 233:478-481.  
30

Depending upon the intended use, the constructs may contain the nucleic acid sequence which encodes the entire LCAT protein, a portion of the LCAT protein, the entire ACAT protein, or a portion of the ACAT protein. For example, where antisense inhibition of

a given LCAT or ACAT protein is desired, the entire sequence is not required. Furthermore, where LCAT or ACAT sequences used in constructs are intended for use as probes, it may be advantageous to prepare constructs containing only a particular portion of a LCAT or ACAT encoding sequence, for example a sequence which is discovered to encode a highly conserved region.

The skilled artisan will recognize that there are various methods for the inhibition of expression of endogenous sequences in a host cell. Such methods include, but are not limited to antisense suppression (Smith, *et al.* (1988) *Nature* 334:724-726), co-suppression (Napoli, *et al.* (1989) *Plant Cell* 2:279-289), ribozymes (PCT Publication WO 97/10328), and combinations of sense and antisense Waterhouse, *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95:13959-13964. Methods for the suppression of endogenous sequences in a host cell typically employ the transcription or transcription and translation of at least a portion of the sequence to be suppressed. Such sequences may be homologous to coding as well as non-coding regions of the endogenous sequence.

Regulatory transcript termination regions may be provided in plant expression constructs of this invention as well. Transcript termination regions may be provided by the DNA sequence encoding the diacylglycerol acyltransferase or a convenient transcription termination region derived from a different gene source, for example, the transcript termination region which is naturally associated with the transcript initiation region. The skilled artisan will recognize that any convenient transcript termination region which is capable of terminating transcription in a plant cell may be employed in the constructs of the present invention.

Alternatively, constructs may be prepared to direct the expression of the LCAT or ACAT sequences directly from the host plant cell plastid. Such constructs and methods are known in the art and are generally described, for example, in Svab, *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530 and Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917 and in U.S. Patent Number 5,693,507.

A plant cell, tissue, organ, or plant into which the recombinant DNA constructs containing the expression constructs have been introduced is considered transformed, transfected, or transgenic. A transgenic or transformed cell or plant also includes progeny of the cell or plant and progeny produced from a breeding program employing such a transgenic plant as a parent in a cross and exhibiting an altered phenotype resulting from the presence of a LCAT nucleic acid sequence.

Plant expression or transcription constructs having a plant LCAT as the DNA sequence of interest for increased or decreased expression thereof may be employed with a wide variety of plant life, particularly, plant life involved in the production of vegetable oils for edible and industrial uses. Most especially preferred are temperate oilseed crops. Plants of interest include, but are not limited to, rapeseed (Canola and High Erucic Acid varieties), sunflower, safflower, cotton, soybean, peanut, coconut and oil palms, and corn. Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences may be required. Importantly, this invention is applicable to dicotyledons and monocotyledons species alike and will be readily applicable to new and/or improved transformation and regulation techniques.

Of particular interest, is the use of plant LCAT and ACAT constructs in plants to produce plants or plant parts, including, but not limited to leaves, stems, roots, reproductive, and seed, with a modified content of lipid and/or sterol esters and to alter the oil production by such plants.

Of particular interest in the present invention, is the use of ACAT genes in conjunction with the LCAT sequences to increase the sterol content of seeds. Thus, overexpression of a nucleic acid sequence encoding an ACAT and LCAT in an oilseed crop may find use in the present invention to increase sterol levels in plant tissues and/or increase oil production.

It is contemplated that the gene sequences may be synthesized, either completely or in part, especially where it is desirable to provide plant-preferred sequences. Thus, all or a portion of the desired structural gene (that portion of the gene which encodes the LCAT or ACAT protein) may be synthesized using codons preferred by a selected host. Host-preferred codons may be determined, for example, from the codons used most frequently in the proteins expressed in a desired host species.

One skilled in the art will readily recognize that antibody preparations, nucleic acid probes (DNA and RNA) and the like may be prepared and used to screen and recover "homologous" or "related" sequences from a variety of plant sources. Homologous sequences are found when there is an identity of sequence, which may be determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions between a known LCAT and a candidate source. Conservative changes, such as Glu/Asp, Val/Ile, Ser/Thr, Arg/Lys and Gln/Asn may also be considered in determining sequence homology. Amino acid sequences are considered homologous by as little as 25% sequence

identity between the two complete mature proteins. (See generally, Doolittle, R.F., *OF URFS and ORFS* (University Science Books, CA, 1986.)

Thus, other LCATs may be obtained from the specific sequences provided herein.

Furthermore, it will be apparent that one can obtain natural and synthetic sequences,

5 including modified amino acid sequences and starting materials for synthetic-protein modeling from the exemplified LCAT and ACAT sequences and from sequences which are obtained through the use of such exemplified sequences. Modified amino acid sequences include sequences which have been mutated, truncated, increased and the like, whether such sequences were partially or wholly synthesized. Sequences which are actually purified from  
10 plant preparations or are identical or encode identical proteins thereto, regardless of the method used to obtain the protein or sequence, are equally considered naturally derived.

For immunological screening, antibodies to the protein can be prepared by injecting rabbits or mice with the purified protein or portion thereof, such methods of preparing antibodies being well known to those in the art. Either monoclonal or polyclonal antibodies  
15 can be produced, although typically polyclonal antibodies are more useful for gene isolation. Western analysis may be conducted to determine that a related protein is present in a crude extract of the desired plant species, as determined by cross-reaction with the antibodies to the encoded proteins. When cross-reactivity is observed, genes encoding the related proteins are isolated by screening expression libraries representing the desired plant species. Expression  
20 libraries can be constructed in a variety of commercially available vectors, including lambda gt11, as described in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

To confirm the activity and specificity of the proteins encoded by the identified nucleic acid sequences as acyltransferase enzymes, *in vitro* assays are performed in insect cell  
25 cultures using baculovirus expression systems. Such baculovirus expression systems are known in the art and are described by Lee, *et al.* U.S. Patent Number 5,348,886, the entirety of which is herein incorporated by reference.

In addition, other expression constructs may be prepared to assay for protein activity utilizing different expression systems. Such expression constructs are transformed into yeast  
30 or prokaryotic host and assayed for acyltransferase activity. Such expression systems are known in the art and are readily available through commercial sources.

The method of transformation in obtaining such transgenic plants is not critical to the instant invention, and various methods of plant transformation are currently available.

Furthermore, as newer methods become available to transform crops, they may also be directly applied hereunder. For example, many plant species naturally susceptible to *Agrobacterium* infection may be successfully transformed via tripartite or binary vector methods of *Agrobacterium* mediated transformation. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders may find use with other modes of transformation. In addition, techniques of microinjection, DNA particle bombardment, and electroporation have been developed which allow for the transformation of various monocot and dicot plant species.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host, viral immunity or the like. Depending upon the number of different host species the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

Non-limiting examples of suitable selection markers include genes that confer resistance to bleomycin, gentamycin, glyphosate, hygromycin, kanamycin, methotrexate, phleomycin, phosphinotricin, spectinomycin, streptomycin, sulfonamide and sulfonylureas. Maliga et al., *Methods in Plant Molecular Biology*, Cold Spring Harbor Laboratory Press, 1995, p. 39. Examples of markers include, but are not limited to, alkaline phosphatase (AP), myc, hemagglutinin (HA),  $\beta$  glucuronidase (GUS), luciferase, and green fluorescent protein (GFP).

Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be introduced into the *Agrobacterium* host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-plasmid containing the T-DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall formation), the latter being permissible, so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed plasmid can give a mixture of normal plant cells and gall.

In some instances where *Agrobacterium* is used as the vehicle for transforming host plant cells, the expression or transcription construct bordered by the T-DNA border region(s)

will be inserted into a broad host range vector capable of replication in *E. coli* and *Agrobacterium*, there being broad host range vectors described in the literature. Commonly used is pRK2 or derivatives thereof. See, for example, Ditta, *et al.*, (*Proc. Nat. Acad. Sci., U.S.A.* (1980) 77:7347-7351) and EPA 0 120 515, which are incorporated herein by  
5 reference. Alternatively, one may insert the sequences to be expressed in plant cells into a vector containing separate replication sequences, one of which stabilizes the vector in *E. coli*, and the other in *Agrobacterium*. See, for example, McBride and Summerfelt (*Plant Mol. Biol.* (1990) 14:269-276), wherein the pRiHRI (Jouanin, *et al.*, *Mol. Gen. Genet.* (1985) 201:370-374) origin of replication is utilized and provides for added stability of the plant  
10 expression vectors in host *Agrobacterium* cells.

Included with the expression construct and the T-DNA can be one or more markers, which allow for selection of transformed *Agrobacterium* and transformed plant cells. A number of markers have been developed for use with plant cells, such as resistance to chloramphenicol, kanamycin, the aminoglycoside G418, hygromycin, or the like. The  
15 particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus  
20 forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils.

Thus, in another aspect of the present invention, methods for modifying the sterol  
25 and/or stanol composition of a host cell. Of particular interest are methods for modifying the sterol and/or stanol composition of a host plant cell. In general the methods involve either increasing the levels of sterol ester compounds as a proportion of the total sterol compounds. The method generally comprises the use of expression constructs to direct the expression of the polynucleotides of the present invention in a host cell.

30 Also provided are methods for reducing the proportion of sterol ester compounds as a percentage of total sterol compounds in a host plant cell. The method generally comprises the use of expression constructs to direct the suppression of endogenous acyltransferase proteins in a host cell.

Of particular interest is the use of expression constructs to modify the levels of sterol compounds in a host plant cell. Most particular, the methods find use in modifying the levels of sterol compounds in seed oils obtained from plant seeds.

Also of interest is the use of expression constructs of the present invention to alter oil production in a host cell and in particular to increase oil production. Of particular interest is the use of expression constructs containing nucleic acid sequences encoding LCAT and/or ACAT polypeptides to transform host plant cells and to use these host cells to regenerate whole plants having increase oil production as compared to the same plant not containing the expression construct.

The oils obtained from transgenic plants having modified sterol compound content find use in a wide variety of applications. Of particular interest in the present invention is the use of the oils containing modified levels of sterol compounds in applications involved in improving human nutrition and cardiovascular health. For example, phytosterols are beneficial for lowering serum cholesterol (Ling, *et al.* (1995) *Life Sciences* 57:195-206).

Cholesterol-lowering compositions comprise the oils and sterol ester compound compositions obtained using the methods of the present invention. Such cholesterol lowering compositions include, but are not limited to foods, food products, processed foods, food ingredients, food additive compositions, or dietary/nutritional supplements that contain oils and/or fats. Non-limiting examples include margarines; butters; shortenings; cooking oils; frying oils; dressings, such as salad dressings; spreads; mayonnaises; and vitamin/mineral supplements. Patent documents relating to such compositions include, U.S. Patents 4,588,717 and 5,244,887, and PCT International Publication Nos. WO 96/38047, WO 97/42830, WO 98/06405, and WO 98/06714. Additional non-limiting examples include toppings; dairy products such as cheese and processed cheese; processed meat; pastas; sauces; cereals; desserts, including frozen and shelf-stable desserts; dips; chips; baked goods; pastries; cookies; snack bars; confections; chocolates; beverages; unextracted seed; and unextracted seed that has been ground, cracked, milled, rolled, extruded, pelleted, defatted, dehydrated, or otherwise processed, but which still contains the oils, etc., disclosed herein.

The cholesterol-lowering compositions can also take the form of pharmaceutical compositions comprising a cholesterol-lowering effective amount of the oils or sterol compound compositions obtained using the methods of the present invention, along with a pharmaceutically acceptable carrier, excipient, or diluent. These pharmaceutical



compositions can be in the form of a liquid or a solid. Liquids can be solutions or suspensions; solids can be in the form of a powder, a granule, a pill, a tablet, a gel, or an extrudate. U.S. Patent 5,270,041 relates to sterol-containing pharmaceutical compositions.

Thus, by expression of the nucleic acid sequences encoding acyltransferase-like sequences of the present invention in a host cell, it is possible to modify the lipid content and/or composition as well as the sterol content and/or composition of the host cell.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

## EXAMPLES

### Example 1: RNA Isolations

Total RNA from the inflorescence and developing seeds of *Arabidopsis thaliana* was isolated for use in construction of complementary (cDNA) libraries. The procedure was an adaptation of the DNA isolation protocol of Webb and Knapp (D.M. Webb and S.J. Knapp, (1990) Plant Molec. Reporter, 8, 180-185). The following description assumes the use of 1g fresh weight of tissue. Frozen seed tissue was powdered by grinding under liquid nitrogen. The powder was added to 10ml REC buffer (50mM Tris-HCl, pH 9, 0.8M NaCl, 10mM EDTA, 0.5% w/v CTAB (cetyltrimethyl-ammonium bromide)) along with 0.2g insoluble polyvinylpyrrolidone, and ground at room temperature. The homogenate was centrifuged for 5 minutes at 12,000 xg to pellet insoluble material. The resulting supernatant fraction was extracted with chloroform, and the top phase was recovered.

The RNA was then precipitated by addition of 1 volume RecP (50mM Tris-HCL pH9, 10mM EDTA and 0.5% (w/v) CTAB) and collected by brief centrifugation as before. The RNA pellet was redissolved in 0.4 ml of 1M NaCl. The RNA pellet was redissolved in water and extracted with phenol/chloroform. Sufficient 3M potassium acetate (pH 5) was added to make the mixture 0.3M in acetate, followed by addition of two volumes of ethanol to precipitate the RNA. After washing with ethanol, this final RNA precipitate was dissolved in water and stored frozen.

Alternatively, total RNA may be obtained using TRIzol reagent (BRL-Lifetechnologies, Gaithersburg, MD) following the manufacturer's protocol. The RNA precipitate was dissolved in water and stored frozen.

## 5 Example 2: Identification of LCAT Sequences

10 Searches were performed on a Silicon Graphics Unix computer using additional Bioaccelerator hardware and GenWeb software supplied by Compugen Ltd. This software and hardware enabled the use of the Smith-Waterman algorithm in searching DNA and protein databases using profiles as queries. The program used to query protein databases was profilesearch. This is a search where the query is not a single sequence but a profile based on a multiple alignment of amino acid or nucleic acid sequences. The profile was used to query a sequence data set, i.e., a sequence database. The profile contained all the pertinent information for scoring each position in a sequence, in effect replacing the "scoring matrix" used for the standard query searches. The program used to query nucleotide databases with a protein profile was tprofilesearch. Tprofilesearch searches nucleic acid databases using an amino acid profile query. As the search is running, sequences in the database are translated to amino acid sequences in six reading frames. The output file for tprofilesearch is identical to the output file for profilesearch except for an additional column that indicates the frame in which the best alignment occurred.

15 The Smith-Waterman algorithm, (Smith and Waterman (1981) *J. Molec. Biol.*, 147:195-197), was used to search for similarities between one sequence from the query and a group of sequences contained in the database.

25 A protein sequence of Lecithin: cholesterol acyltransferase from human (McLean J, *et al.* (1986) *Nucleic Acids Res.* 14(23):9397-406 SEQ ID NO:1)) was used to search the NCBI non-redundant protein database using BLAST. Three sequences were identified from *Arabidopsis*, GenBank accessions AC004557 (referred to herein as LCAT1, SEQ ID NO:2), AC003027 (referred to herein as LCAT2, SEQ ID NO:4), and AL024486 (referred to herein as LCAT3, SEQ ID NO:6). The deduced amino acid sequences are provided in SEQ ID NOs: 30 3, 5, and 7, respectively.

The profile generated from the queries using PSI-BLAST was excised from the hyper text markup language (html) file. The worldwide web (www)/html interface to psiblast at ncbi stores the current generated profile matrix in a hidden field in the html file that is

returned after each iteration of psiblast. However, this matrix has been encoded into string62 (s62) format for ease of transport through html. String62 format is a simple conversion of the values of the matrix into html legal ascii characters.

The encoded matrix width (x axis) is 26 characters, and comprise the consensus characters, the probabilities of each amino acid in the order A,B,C,D,E,F,G,H,I,K,L,M,N, P,Q,R,S,T,V,W,X,Y,Z (where B represents D and N, and Z represents Q and E, and X represents any amino acid), gap creation value, and gap extension value.

The length (y axis) of the matrix corresponds to the length of the sequences identified by PSI-BLAST. The order of the amino acids corresponds to the conserved amino acid sequence of the sequences identified using PSI-BLAST, with the N-terminal end at the top of the matrix. The probabilities of other amino acids at that position are represented for each amino acid along the x axis, below the respective single letter amino acid abbreviation.

Thus, each row of the profile consists of the highest scoring (consensus) amino acid, followed by the scores for each possible amino acid at that position in sequence matrix, the score for opening a gap at that position, and the score for continuing a gap at that position.

The string62 file is converted back into a profile for use in subsequent searches. The gap open field is set to 11 and the gap extension field is set to 1 along the x axis. The gap creation and gap extension values are known, based on the settings given to the PSI-BLAST algorithm. The matrix is exported to the standard GCG profile form. This format can be read by GenWeb.

The algorithm used to convert the string62 formatted file to the matrix is outlined in Table 1.

**Table 1**

1. if encoded character z then the value is blast score min
2. if encoded character Z then the value is blast score max
3. else if the encoded character is uppercase then its value is (64-(ascii # of char))
4. else if the encoded character is a digit the value is ((ascii # of char)-48)
5. else if the encoded character is not uppercase then the value is ((ascii # of char) - 87)
6. ALL B positions are set to min of D and N amino acids at that row in sequence matrix
7. ALL Z positions are set to min of Q and E amino acids at that row in sequence matrix
8. ALL X positions are set to min of all amino acids at that row in sequence matrix
9. kBLAST\_SCORE\_MAX=999;
10. kBLAST\_SCORE\_MIN=-999;
11. all gap opens are set to 11
12. all gap lens are set to 1

The protein sequences of LCAT1, LCAT2, and LCAT3 as well as the PSI-BLAST profile were used to search public and proprietary databases for additional LCAT sequences. Two EST sequences were identified which appear to be identical to LCAT1 and LCAT3, respectively. One additional *Arabidopsis* sequence was identified from the proprietary databases, LCAT4 (SEQ ID NO:8). The deduced protein sequence of LCAT4 is provided in SEQ ID NO:9. Two additional genomic sequences were identified using the PSI-BLAST profile from libraries of *Arabidopsis* ecotypes Columbia and Landsberg, LCAT7 (SEQ ID NO:10) and LCAT8 (SEQ ID NO:11). The LCAT7 sequence was present in both the Columbia and Landsberg genomic libraries, while the LCAT8 sequence was only present in the Columbia library.

An open reading frame was predicted from the genomic sequence of LCAT7 in the *Arabidopsis* public database and this sequence was called MSH12 (referred to herein as LCAT5, SEQ ID NO: 73). The deduced protein sequence of LCAT5 is provided in SEQ ID NO: 74.

The PSI-BLAST profile and the LCAT sequences were used to query the public yeast database and proprietary libraries containing corn and soy EST sequences. The yeast genome contains only one gene, *LRO1* (LCAT Related Open reading frame, YNR008W, Figure 1) with distinct similarity to the human LCAT. The DNA sequence of *LRO1* is provided in SEQ

ID NO: 75 and the protein sequence is provided in SEQ ID NO: 76. Seven EST sequences were identified from soybean libraries as being LCAT sequences. Two sequences from soy (SEQ ID NOs: 12 and 13) are most closely related to the *Arabidopsis* LCAT1 sequence, a single sequence was identified as being most closely related to LCAT2 (SEQ ID NO:14), three were closely related to LCAT3 (SEQ ID NOs: 15-17), and an additional single sequence was identified (SEQ ID NO:18). A total of 11 corn EST sequences were identified as being related to the *Arabidopsis* LCAT sequences. Two corn EST sequences (SEQ ID NOs: 19 and 20) were most closely related to LCAT1, two sequences were identified as closely related to LCAT2 (SEQ ID NOs: 21 and 22), four corn EST sequences were identified as closely related to LCAT3 (SEQ ID NOs: 23-26), and an additional three corn EST sequences were also identified (SEQ ID NOs: 27-29).

### Example 3: Identification of ACAT Sequences

Since plant ACATs are unknown in the art, searches were performed to identify known and related ACAT sequences from mammalian sources from public databases. These sequences were then used to search public and proprietary EST databases to identify plant ACAT-like sequences.

A public database containing mouse Expressed Sequence Tag (EST) sequences (dbEST) was searched for ACAT-like sequences. The search identified two sequences (SEQ ID NO 30 and 31) which were related (approximately 20% identical), but divergent, to known ACAT sequences.

In order to identify ACAT-like sequences from other organisms, the two mouse ACAT sequences were used to search public and proprietary databases containing EST sequences from human and rat tissues. Results of the search identified several sequences from the human database and from the rat database which were closely related to the mouse sequences. The human and rat ACAT-like EST sequences were assembled, using the GCG assembly program, to construct a complete inferred cDNA sequence by identifying overlapping sequences (SEQ ID NOs: 32 and 33, respectively).

The protein sequence of the human ACAT-like sequence was aligned with known ACAT sequences from human (Chang, *et al.* (1993) *J. Biol. Chem.* 268:20747-20755, SEQ ID NO:34), mouse (Uelmen, *et al.* (1995) *J. Biol. Chem.* 270:26192-26201 SEQ ID NO:35) and yeast (Yu, *et al.* (1996) *J. Biol. Chem.* 271:24157-24163, SEQ ID NO:36 and Yang, *et al.*

(1996) *Science* 272:1353-1356, SEQ ID NO:37) using MacVector (Oxford Molecular, Inc.). Results of the alignment demonstrated that the sequence was related to the known sequences, however the related sequence was only about 25% similar to the known sequences.

The protein sequence of the human sterol O-acyltransferase (ACAT, Acyl CoA:Cholesterol acyltransferase, Accession number A48026) related sequence was used to search protein and nucleic acid Genbank databases. A single plant homologue was identified in the public *Arabidopsis* EST database (Accession A042298, SEQ ID NO:38). The protein sequence (SEQ ID NO:39) was translated from the EST sequence, and was found to contain a peptide sequence conserved in both mammalian and yeast ACATs (Chang *et al.*, (1997) *Ann. Rev. Biochem.*, 66:613-638).

To obtain the entire coding region corresponding to the *Arabidopsis* ACAT-like EST, synthetic oligo-nucleotide primers were designed to amplify the 5' and 3' ends of partial cDNA clones containing ACAT-like sequences. Primers were designed according to the *Arabidopsis* ACAT-like EST sequence and were used in Rapid Amplification of cDNA Ends (RACE) reactions (Frohman *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:8998-9002).

Primers were designed (5'-TGCAAATTGACGAGCACACCAACCCCTTC-3' (SEQ ID NO:40) and 5'-AAGGATGCTTTGAGTTCCTGACAATAGG-3' (SEQ ID NO:41)) to amplify the 5' end from the *Arabidopsis* ACAT EST sequence. Amplification of flanking sequences from cDNA clones were performed using the Marathon cDNA Amplification kit (Clontech, CA).

The sequence derived from the 5'-RACE amplification was used to search proprietary *Arabidopsis* EST libraries. A single EST accession, LIB25-088-C7 (SEQ ID NO:42), was identified which contained a sequence identical to the 5'-RACE sequence. Furthermore, LIB25-088-C7 was found to contain the complete putative coding sequence for the *Arabidopsis* ACAT-like product.

The nucleic acid as well as the putative translation product sequences of A042298 were used to search public and proprietary databases. Four EST sequences were identified in both soybean (SEQ ID NOs:43-46) and maize (SEQ ID NOs:47-50) proprietary databases, and a single ACAT-like sequence was identified from *Mortierella alpina* EST sequences (SEQ ID NO:51).

Sequence alignments between ACAT sequences from several different sources were compared to identify the similarity between the sequences. Nucleotide sequences from

known human and mouse ACATs, as well as nucleotide sequences from known yeast ACATs were compared to the ACAT-like EST sequences from human and *Arabidopsis*.

Analysis of the sequence alignments revealed several classes of ACATs based on sequence similarity. The known human and mouse ACATs, being 88% similar in the nucleotide sequence, formed one class of ACATs. Another class of ACATs included the yeast ACATs which are less than 20% similar to the known human and mouse class ACATs.

The final class of ACATs included the *Arabidopsis* and human sequences disclosed in the present invention. This class is approximately 22% similar to the known human and mouse ACAT class and approximately 23% similar to the yeast class of ACATs. Thus, the ACAT sequences disclosed in the present invention represent a novel class of ACAT enzymes. Partial mouse sequences of this class are also provided.

#### **Example 4: Expression Construct Preparation**

Constructs were prepared to direct expression of the LCAT1, LCAT2, LCAT3, LCAT4, LCAT5 and the yeast LRO1 sequences in plants and cultured insect cells. The entire coding region of each LCAT was amplified from the appropriate EST clone or an *Arabidopsis* genomic cDNA library using the following oligonucleotide primers in a polymerase chain reactions (PCR). The LCAT1 coding sequence was amplified from the EST clone Lib25-082-Q1-E1-G4 using the primers 5'-GGATCCGCGGCCGCACAATGAAAAAATATCTTCACATTATTCGG-3' (SEQ ID NO:52) and 5'-GGATCCCCTGCAGGTCATTGACGGCATTAAACATTGG-3' (SEQ ID NO:53). The LCAT2 coding sequence was amplified from an *Arabidopsis* genomic cDNA library using the synthetic oligo nucleotide primers 5'-GGATCCGCGGCCGCACAATGGGAGCGAATTCGAAATCAGTAACG-3' (SEQ ID NO:54) and 5'-GGATCCCCTGCAGGTTAATACCCACTTTTATCAAGCTCCC-3' (SEQ ID NO:55). The LCAT3 coding sequence was amplified from the EST clone LIB22-004-Q1-E1-B4 using the synthetic oligo nucleotide primers 5'-GGATCCGCGGCCGCACAATGTCTCTATTACTGGAA GAGATC-3' (SEQ ID NO:56) and 5'-GGATCCCCTGCAGGTTATGCATC AACAGAGACACTTACAGC-3' (SEQ ID NO:57). The LCAT4 coding sequence was amplified from the EST clone LIB23-007-Q1-E1-B5 using the synthetic oligo nucleotide primers

5'-GGATCCGCGGCCGCACAATGGGCTGGATTCCGTGTCCGTGC-3' (SEQ ID NO:58)  
and 5'-GGATCCCCTGCAGGTTAACCAGAATCAACTACTTTGTG-3' (SEQ ID NO:59).  
The LCAT5 coding sequence was amplified from LIB23-053-Q1-E1-E3 using the synthetic  
oligo nucleotide primers

- 5 5'-GGATCCGCGGCCGCACAATGCCCCTTATTCATCGG-3' (SEQ ID NO:77) and 5'-  
GGATCCCCTGCAGGTCACAGCTTCAGGTCAATACG-3' (SEQ ID NO:78).

The yeast LROI coding sequence was amplified from genomic yeast DNA using the  
synthetic oligo nucleotide primers

- 5'GGATCCGCGGCCGCACAATGGGCACACTGTTTCGAAG3' (SEQ ID NO:79)  
10 and 5'GGATCCCCTGCAGGTTACATTGGGAAGGGCATCTGAG3' (SEQ ID NO:80).

The entire coding region of the *Arabidopsis* ACAT sequence (SEQ ID NO: 42) was  
amplified from the EST clone LIB25-088-C7 using oligonucleotide primers

- 5'-TCGACCTGCAGGAAGCTTAGAAATGGCGATTTTGGATTTC-3' (SEQ ID NO: 60)  
and 5'-GGATCCGCGGCCGCTCATGACATCGATCCTTTTCGG-3' (SEQ ID NO: 61) in a  
15 polymerase chain reaction (PCR).

Each resulting PCR product was subcloned into pCR2.1Topo (Invitrogen) and labeled  
pCGN9964 (LCAT1), pCGN9985 (LCAT2), pCGN9965 (LCAT3), pCGN9995 (LCAT4),  
pCGN10964 (LCAT5), pCGN10963 (*LROI*), and pCGN8626 (ACAT). Double stranded  
DNA sequence was obtained to verify that no errors were introduced by the PCR  
20 amplification.

#### 4A. Baculovirus Expression Constructs

Constructs are prepared to direct the expression of the *Arabidopsis* LCAT and yeast  
25 LCAT sequences in cultured insect cells. The entire coding region of the LCAT proteins was  
removed from the respective constructs by digestion with *NotI* and *Sse8387I*, followed by gel  
electrophoresis and gel purification. The fragments containing the LCAT coding sequences  
were cloned into *NotI* and *PstI* digested baculovirus expression vector pFastBac1 (Gibco-  
BRL, Gaithersburg, MD). The resulting baculovirus expression constructs were referred to as  
30 pCGN9992 (LCAT1), pCGN9993 (LCAT2), pCGN9994 (LCAT3), pCGN10900 (LCAT4),  
pCGN10967 (LCAT5), and pCGN10962 (*LROI*).



## 4B. Plant Expression Construct Preparation

A plasmid containing the napin cassette derived from pCGN3223 (described in U.S. Patent No. 5,639,790, the entirety of which is incorporated herein by reference) was modified to make it more useful for cloning large DNA fragments containing multiple restriction sites, and to allow the cloning of multiple napin fusion genes into plant binary transformation vectors. An adapter comprised of the self annealed oligonucleotide of sequence 5'-CGCGATTAAATGGCGCGCCCTGCAGGCGGCCGCTGCAGGGCGCGCCATTAAAT-3' (SEQ ID NO:62) was ligated into the cloning vector pBC SK+ (Stratagene) after digestion with the restriction endonuclease BssHII to construct vector pCGN7765. Plasmids pCGN3223 and pCGN7765 were digested with NotI and ligated together. The resultant vector, pCGN7770, contained the pCGN7765 backbone with the napin seed specific expression cassette from pCGN3223.

The cloning cassette, pCGN7787, contained essentially the same regulatory elements as pCGN7770, with the exception of the napin regulatory regions of pCGN7770 have been replaced with the double CAMV 35S promoter and the tml polyadenylation and transcriptional termination region.

A binary vector for plant transformation, pCGN5139, was constructed from pCGN1558 (McBride and Summerfelt, (1990) Plant Molecular Biology, 14:269-276). In pCGN5139, the polylinker of pCGN1558 was replaced as a HindIII/Asp718 fragment with a polylinker containing unique restriction endonuclease sites, AscI, PacI, XbaI, SwaI, BamHI, and NotI. The Asp718 and HindIII restriction endonuclease sites are retained in pCGN5139.

A series of turbo binary vectors was constructed to allow for the rapid cloning of DNA sequences into binary vectors containing transcriptional initiation regions (promoters) and transcriptional termination regions.

The plasmid pCGN8618 was constructed by ligating oligonucleotides 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:63) and 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC-3' (SEQ ID NO:64) into SalI/XhoI-digested pCGN7770. A fragment containing the napin promoter, polylinker and napin 3' region was excised from pCGN8618 by digestion with Asp718I; the fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5'

overhangs with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated

5 pCGN8622.

The plasmid pCGN8619 was constructed by ligating oligonucleotides 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC-3' (SEQ ID NO:65) and 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:66) into SalI/XhoI-digested pCGN7770. A fragment containing the napin promoter, polylinker and  
10 napin 3' region was removed from pCGN8619 by digestion with Asp718I; the fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest  
15 to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8623.

The plasmid pCGN8620 was constructed by ligating oligonucleotides 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGGAGCT -3' (SEQ ID NO:67) and  
20 5'-CCTGCAGGAAGCTTGCGGCCGCGGATCC-3' (SEQ ID NO:68) into SalI/SacI-digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region was removed from pCGN8620 by complete digestion with Asp718I and partial digestion with NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-  
25 ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8624.

30 The plasmid pCGN8621 was constructed by ligating oligonucleotides 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCCAGCT-3' (SEQ ID NO:69) and 5'-GGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:70) into SalI/SacI-digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region

was removed from pCGN8621 by complete digestion with Asp718I and partial digestion with NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8625.

The plasmid construct pCGN8640 is a modification of pCGN8624 described above. A 938bp PstI fragment isolated from transposon Tn7 which encodes bacterial spectinomycin and streptomycin resistance (Fling et al. (1985), *Nucleic Acids Research* 13(19):7095-7106), a determinant for *E. coli* and *Agrobacterium* selection, was blunt ended with Pfu polymerase. The blunt ended fragment was ligated into pCGN8624 that had been digested with SpeI and blunt ended with Pfu polymerase. The region containing the PstI fragment was sequenced to confirm both the insert orientation and the integrity of cloning junctions.

The spectinomycin resistance marker was introduced into pCGN8622 and pCGN8623 as follows. A 7.7 Kbp AvrII-SnaBI fragment from pCGN8640 was ligated to a 10.9 Kbp AvrII-SnaBI fragment from pCGN8623 or pCGN8622, described above. The resulting plasmids were pCGN8641 and pCGN8643, respectively.

The plasmid pCGN8644 was constructed by ligating oligonucleotides 5'-GATCACCTGCAGGAAGCTTGC GGCCGCGGATCCAATGCA-3' (SEQ ID NO:71) and 5'-TTGGATCCGCGGCCGCAAGCTTCCTGCAGGT-3' (SEQ ID NO:72) into BamHI-PstI digested pCGN8640.

#### 4C. Plant LCAT Expression Construct Preparation

The coding sequence of LCAT1 was cloned from pCGN9964 as a *NotI*/ *Sse8387I* fragment into pCGN8640, pCGN8641, pCGN8643, and pCGN8644 to create the expression constructs pCGN9960, pCGN9961, pCGN9962, and pCGN9963, respectively. The construct pCGN9960 was designed to express the LCAT1 coding sequence in the sense orientation from the constitutive promoter CaMV 35S. The construct pCGN9961 was designed to express the LCAT1 coding sequence in the antisense orientation from the napin promoter. The construct pCGN9962 was designed to express the LCAT1 coding sequence in the sense

orientation from the napin promoter. The construct pCGN9963 was designed to express the LCAT1 coding sequence in the antisense orientation from the constitutive promoter CaMV 35S.

The coding sequence of LCAT2 was cloned from pCGN9985 as a *NotI*/*Sse8387I* fragment into pCGN8640, pCGN8641, pCGN8643, and pCGN8644 to create the expression constructs pCGN9981, pCGN9982, pCGN9983, and pCGN9984, respectively. The construct pCGN9981 was designed to express the LCAT2 coding sequence in the sense orientation from the constitutive promoter CaMV 35S. The construct pCGN9982 was designed to express the LCAT2 coding sequence in the antisense orientation from the napin promoter.

The construct pCGN9983 was designed to express the LCAT2 coding sequence in the sense orientation from the napin promoter. The construct pCGN9984 was designed to express the LCAT2 coding sequence in the antisense orientation from the constitutive promoter CaMV 35S.

The coding sequence of LCAT3 was cloned from pCGN9965 as a *NotI*/*Sse8387I* fragment into pCGN8640, pCGN8641, pCGN8643, and pCGN8644 to create the expression constructs pCGN9966, pCGN9967, pCGN9968, and pCGN9969, respectively. The construct pCGN9966 was designed to express the LCAT3 coding sequence in the sense orientation from the constitutive promoter CaMV 35S. The construct pCGN9967 was designed to express the LCAT3 coding sequence in the antisense orientation from the napin promoter.

The construct pCGN9968 was designed to express the LCAT3 coding sequence in the sense orientation from the napin promoter. The construct pCGN9969 was designed to express the LCAT3 coding sequence in the antisense orientation from the constitutive promoter CaMV 35S.

The coding sequence of LCAT4 was cloned from pCGN9995 as a *NotI*/*Sse8387I* fragment into pCGN8640, pCGN8641, pCGN8643, and pCGN8644 to create the expression constructs pCGN9996, pCGN9997, pCGN9998, and pCGN9999, respectively. The construct pCGN9996 was designed to express the LCAT4 coding sequence in the sense orientation from the constitutive promoter CaMV 35S. The construct pCGN9997 was designed to express the LCAT4 coding sequence in the antisense orientation from the napin promoter.

The construct pCGN9998 was designed to express the LCAT4 coding sequence in the sense orientation from the napin promoter. The construct pCGN9999 was designed to express the LCAT4 coding sequence in the antisense orientation from the constitutive promoter CaMV 35S.

The coding sequence of *LCAT5* was cloned from pCGN10964 as a *NotI*/*Sse8387I* fragment into pCGN9977 and pCGN9979, to create the expression constructs pCGN10965, and pCGN10966, respectively. The construct pCGN10965 was designed to express the *LCAT5* coding sequence in the sense orientation from the constitutive promoter CaMV 35S.

5 The construct pCGN10966 was designed to express the *LCAT5* coding sequence in the sense orientation from the napin promoter.

The coding sequence of *LRO1* was cloned from pCGN10963 as a *NotI*/*Sse8387I* fragment into pCGN9977 and pCGN9979, to create the expression constructs pCGN10960, and pCGN10961, respectively. The construct pCGN10960 was designed to express the  
10 *LRO1* coding sequence in the sense orientation from the constitutive promoter CaMV 35S. The construct pCGN10961 was designed to express the *LRO1* coding sequence in the sense orientation from the napin promoter.

#### 4D. Plant ACAT Expression Construct Preparation

15

A fragment containing the *Arabidopsis* ACAT-like coding region was removed from pCGN8626 by digestion with *Sse8387I* and *Not I*. The fragment containing the ACAT-like sequence was ligated into *PstI*-*Not I* digested pCGN8622. The resulting plasmid was designated pCGN8627. DNA sequence analysis confirmed the integrity of the cloning  
20 junctions.

A fragment containing the *Arabidopsis* ACAT-like coding region (SEQ ID NO: 42) was removed from pCGN8626 by digestion with *Sse8387I* and *Not I*. The fragment was ligated into *PstI*-*Not I* digested pCGN8623. The resulting plasmid was designated pCGN8628. DNA sequence analysis confirmed the integrity of the cloning junctions.

25

A fragment containing the *Arabidopsis* ACAT-like coding region was removed from pCGN8626 by digestion with *Sse8387* and *Not I*. The fragment was ligated into *PstI*-*Not I* digested pCGN8624. The resulting plasmid was designated pCGN8629. DNA sequence analysis confirmed the integrity of the cloning junctions.

30

A fragment containing the *Arabidopsis* ACAT-like coding region was removed from pCGN8626 by digestion with *Sse8387* and *Not I*. The fragment was ligated into *PstI*-*Not I* digested pCGN8625. The resulting plasmid was designated pCGN8630. DNA sequence analysis confirmed the integrity of the cloning junctions.

An additional expression construct for the suppression of endogenous ACAT-like activity was also prepared. The construct pCGN8660 was constructed by cloning approximately 1 Kb of the *Arabidopsis* ACAT-like coding region from pCGN8626 in the sense orientation, and the full-length *Arabidopsis* ACAT-like coding region in the antisense orientation under the regulatory control of the napin transcription initiation sequence.

For expression of the rat ACAT-like sequence in plants, the NotI-Sse8387I fragment of pCGN8592 was cloned into NotI-PstI digested binary vectors pCGN8621, pCGN8622, and pCGN8624 to yield plasmids, pCGN 9700, pCGN9701, and pCGN9702, respectively. Plasmid pCGN9700 expresses a sense transcript of the rat ACAT-like cDNA under control of a napin promoter, plasmid pCGN9701 expresses an antisense transcript of the rat ACAT-like cDNA under control of a napin promoter, and plasmid pCGN9702 expresses a sense transcript of the rat ACAT-like cDNA under control of a double 35S promoter. Plasmids pCGN 9700, pCGN9701, and pCGN9702 were introduced in *Agrobacterium tumefaciens* EHA101.

Constructs were prepared to direct the expression of the rat ACAT-like sequence in the seed embryo of soybean and the endosperm of corn. For expression of the rat ACAT-like DNA sequence in soybean, a 1.5 kb NotI/Sse8387I fragment from pCGN8592 containing the coding sequence of the rat ACAT-like sequence was blunt ended using Mung bean nuclease, and ligated into the SmaI site of the turbo 7S binary/cloning vector pCGN8809 to create the vector pCGN8817 for transformation into soybean by particle bombardment. The vector pCGN8817 contained the operably linked components of the promoter region of the soybean  $\alpha'$  subunit of  $\beta$ -conglycinin (7S promoter, (Chen *et al.*, (1986), *Proc. Natl. Acad. Sci.*, 83:8560-8564), the DNA sequence coding for the entire rat ACAT-like protein, and the transcriptional termination region of pea RuBisCo small subunit, referred to as E9 3' (Coruzzi, *et al.* (1984) *EMBO J.* 3:1671-1679 and Morelli, *et al.* (1985) *Nature* 315:200-204). This construct further contained sequences for the selection of positive transformed plants by screening for resistance to glyphosate using the CP4 EPSPS (U.S. Patent 5,633,435) expressed under the control of the figwort mosaic virus (FMV) promoter (U.S. Patent Number 5,378,619) and the transcriptional termination region of E9.

For expression of the rat ACAT-like sequence in the corn endosperm, a 1.5 kb NotI/Sse8387I fragment from pCGN8592 containing the coding sequence of the rat ACAT-like sequence was blunt ended using Mung bean nuclease, and ligated into the BamHI site of the rice pGt1 expression cassette pCGN8592 for expression from the pGt1 promoter (Leisy,

D.J. et al., Plant Mol. Biol. 14 (1989) 41-50) and the HSP70 intron sequence (U.S. Patent Number 5,593,874). This cassette also included the transcriptional termination region downstream of the cloning site of nopaline synthase, *nos* 3' (Depicker et al., *J. Molec. Appl. Genet.* (1982) 1: 562-573). A 7.5 kb fragment containing the pGt1 promoter, the DNA sequence encoding the rat ACAT-like protein, and the *nos* transcriptional termination sequence was cloned into the binary vector pCGN8816 to create the vector pCGN8818 for transformation into corn. This construct also contained sequences for the selection of positive transformants with kanamycin using the kanamycin resistance gene from Tn5 bacteria under the control of the CAMV 35S promoter and tml transcriptional termination regions.

#### Example 5: Expression in Insect Cell Culture

A baculovirus expression system was used to express the LCAT cDNAs in cultured insect cells.

The baculovirus expression constructs pCGN9992, pCGN9993, pCGN9994, pCGN10900, pCGN10962, and pCGN10967 were transformed and expressed using the BAC-to-BAC Baculovirus Expression System (Gibco-BRL, Gaithersburg, MD) according to the manufacturer's directions.

The transformed insect cells were used to assay for acyltransferase activities using methods known in the art (see Example 8).

#### Example 6: Plant Transformation

A variety of methods have been developed to insert a DNA sequence of interest into the genome of a plant host to obtain the transcription or transcription and translation of the sequence to effect phenotypic changes. Transgenic plants were obtained by *Agrobacterium*-mediated transformation as described by Radke et al. (*Theor. Appl. Genet.* (1988) 75:685-694; *Plant Cell Reports* (1992) 11:499-505). Alternatively, microprojectile bombardment methods, such as described by Klein et al. (*Bio/Technology* 10:286-291) may also be used to obtain nuclear transformed plants. Other plant species may be similarly transformed using related techniques.

The plant binary constructs described above were used in plant transformation to direct the expression of the sterol acyltransferases in plant tissues. Binary vector constructs were transformed into strain EHA101 *Agrobacterium* cells (Hood et al., *J. Bacteriol* (1986) 168:1291-

1301), by the method of Holsters *et al.* (*Mol. Gen. Genet.* (1978) 163:181-187). Transgenic *Arabidopsis thaliana* plants were obtained by *Agrobacterium*-mediated transformation as described by Valverkens *et al.*, (*Proc. Nat. Acad. Sci.* (1988) 85:5536-5540), Bent *et al.* ((1994), *Science* 265:1856-1860), and Bechtold *et al.* ((1993), *C. R. Acad. Sci., Life Sciences* 316:1194-1199).

#### EXAMPLE 7: Plant Assays for Modified Sterol Content/Profile

##### 7a: NMR of T2 seed

Seed from plants expressing LCAT 1 through 4 under the control of the napin promoter were analyzed by NMR. Arabidopsis seeds from transgenic plants were placed directly into wide-mouth MAS NMR sample tubes.

High-resolution spectra were measured at 11.7 T (1H=500 MHz, 13C=125 MHz) using Varian NMR Instruments (Palo Alto, CA) Inova™ NMR spectrometers equipped with carbon-observe MAS Nanoprobes™. The 13C spectra were acquired without a field-frequency lock at ambient temperature (approx. 21-22°C) for 14 hours using the following conditions: spectral width = 29.996 kHz, acquisition time = 2.185 seconds, p/2 pulse (3.8 ms) with no relaxation delay, 1H g B2 = 2.5 kHz with Waltz decoupling. Data processing conditions were typically: digital resolution = 0.11 Hz, 0.3 to 1.5 Hz line broadening and time-reversed linear prediction of the first three data points. Chemical shifts were referenced by adding neat tetramethylsilane (TMS) to Arabidopsis seeds and using the resulting referencing parameters for subsequent spectra. The 13C resolution was 2-3 Hz for the most narrow seed resonances. Spectral resolution was independent of MAS spinning speeds (1.5-3.5 kHz) and data were typically obtained with 1.5 kHz spinning speeds. Spinning sidebands were approx. 1% of the main resonance. Phytosterol 13C assignments were based on model samples composed of triolein,  $\beta$ -sitosterol and cholesterol oleate. Triacylglycerol 13C assignments were made from comparison with literature assignments or with shifts computed from a 13C NMR database (Advanced Chemical Development, Inc., version 3.50, Toronto Canada).

The results of these analyses are displayed in Figure 2 and show that there was a trend of an approximately 2 fold increase of phytosterols in the seeds derived from plant line 5 expressing the LCAT 4 gene (pCGN9998) under the control of the napin promoter. During the course of this analysis it was also noted that the average oil content of seed from plants



5

4

10

15

20

25

30

## EXAMPLE 8: Baculovirus Insect Cell Culture for Sterol Esterification Activity

Baculovirus expression construct pCGN9992, pCGN9993, pCGN9994 and pCGN10900 (see Example 4) were transformed and expressed using the BAC-TOBAC  
5 Baculovirus Expression System (Gibco-BRL, Gaithersburg, MD) according to the manufacturer's instructions except harvesting of recombinant viruses was done 5 days post-transfection. The supernatant from the transfection mixture was used for generating virus stock which in turn was used for infecting Sf9 cells used in the assay.

The transformed cells were assayed for lecithin:sterol acyltransferase activities using  
10 the method described herein. Insect cells were centrifuged and the resulting cell pellet was either used immediately or stored at -70 C for later analysis. Cells were resuspended in Medium A (100 mM Tricine/NaOH, pH 7.8, 10% (w/v) glycerol, 280 mM NaCl with : 0.1  $\mu$ M Aprotinin, 1  $\mu$ M Leupeptin, and 100  $\mu$ M Pefabloc (all from Boehringer Mannheim, Germany) and lysed by sonication (2 x 10 sec). Cell walls and other debris were pelleted by  
15 centrifugation (14,000 x g , 10 min, 4°C). The supernatant was transferred to a new vial and membranes pelleted by centrifugation (100,000 x g, Ti 70.1 rotor, 46,000 rpm for 1 hour at 4°C). Total membranes were resuspended in Medium A. Lecithin:sterol acyltransferase activity was assayed in a 0.1 ml reaction mixture containing 100 mM Tris/HCl, pH 7, 28 mM NaCl, 0.03% Triton X-100, 0.1 mM sitosterol, 20  $\mu$ M 1,2-[<sup>14</sup>C]-palmitoyl-phosphatidyl  
20 choline (246420 dpm/nmole), and 0.05-20 mg of membrane protein. After 15 minutes at 30 °C, the reaction was terminated by addition of a 0.5 ml solution of methylene chloride:methanol (4:1, v/v ) containing 100  $\mu$ g cholesterol and cholesterol ester as cold carriers. A portion (0.1 ml) of the bottom organic layer was removed and evaporated under nitrogen gas. The concentrated extract was resuspended in 30  $\mu$ l of hexane and spotted onto a  
25 silica gel-G thin layer chromatographic plate. The plate was migrated in hexane:diethyl ether:acetic acid (80:20:1) to the top, then air dried. Radioactivity was determined by exposure to a Low Energy Phosphor-imaging Screen. Following exposure, the screen was read on a phosphorimager.

The LCAT 4 protein from pCGN10900 in baculovirus membranes showed a  
30 radioactive spot in the region of the TLC plate where cholesterol ester migrates indicating that LCAT 4 has the ability to catalyze the transfer of an acyl group from lecithin (PC) to sitosterol to make a sitosterol ester.

**EXAMPLE 9: Plant Assay for Modified Lipid Content**

Nir (near infrared spectroscopy spectral scanning) can be used to determine the total oil content of Arabidopsis seeds in a non-destructive way provided that a spectral calibration curve has been developed and validated for seed oil content. A seed oil spectral calibration curve was developed using seed samples from 85 Arabidopsis plants. Seed was cleaned and scanned using a Foss NIR model 6500 (Foss-Nirs Systems, Inc.). Approximately 50 to 100 milligrams of whole seeds, per sample, were packed in a mini sample ring cup with quartz lens [ IH-0307 ] consisting a mini-insert [ IH-0337 ] and scanned in reflectance mode to obtain the spectral data. The seed samples were then ground, using a mortar and pestle, to achieve a fine homogeneous meal. The ground samples were measured for oil using an accelerated solvent extractor (ASE).

Measurement for the total oil content was performed on the Dionex Accelerated Solvent Extractor (ASE). Approximately 500 mg of clean ground seed was weighed to the nearest 0.1 mg onto a 9 x 9 cm weigh boat. An equal amount of diatomaceous earth was added using a top-loading balance accurate to the nearest 0.01 g. The ground seed sample and the diatomaceous earth were thoroughly mixed until a homogeneous texture was achieved. The sample was loaded on to the instrument and oil extraction was achieved using hexane under validated laboratory protocols. Standard Rapeseed samples were obtained from the Community Bureau of Reference (BCR). The ASE extraction method was validated using the BCR reference standards. A total percent oil recovery of 99% to 100% was achieved. "As-is" oil content was calculated to the nearest 0.01 mass percentage using the formula:

$$\text{Oil Content} = 100\% \times (\text{vial plus extracted oil wt} - \text{initial vial wt}) / (\text{sample wt})$$

The analytical data generated by ASE were used to perform spectral calibrations. Nir calibration equations were generated using the built-in statistical package within the NirSystems winisi software. The spectral calibration portion of the software is capable of calibration and self-validation. From a total of 85 samples, 57 samples were used to generate the total percent oil calibration. The remaining samples were used to validate the oil calibrations. Optimized smoothing, derivative size, and mathematical treatment (modified partial least square) was utilized to generate the calibration. The samples that were not used

in building respective calibrations were used as a validation set. Statistical tools such as correlation coefficient (  $R$  ), coefficient of determination ( $R^2$ ), standard error of prediction ( SEP ), and the standard error of prediction corrected for bias (SEPC) were used to evaluate the calibration equations.

5 T2 seeds from plants that had been transformed with the LCAT genes were cleaned and scanned using a Foss NIR model 6500 (Foss-Nirs Systems, Inc.). Approximately 50 to 100 milligrams of whole seeds, per sample, were packed in a mini sample ring cup with quartz lens [ IH-0307 ] consisting a mini-insert [ IH-0337 ] and scanned in reflectance mode to obtain the spectral data. Oil percentage in each seed sample was determined using the seed  
10 oil spectral calibration curve detailed above.

The results of these analyses are found in Figure 5 and Table 2 and show that there was a significant increase in the oil level in seed from T2 plants expressing the LCAT2 gene. This increase in oil was seen in plants when LCAT2 was driven by either the 35S constitutive promoter or the seed-specific napin promoter. These results show that overexpression of a  
15 nucleic acid sequence encoding a lecithin:cholesterol acyltransferase-like polypeptide can increase seed oil production in plants.

09651651-083000

[illegible]

	Construct number	Seed Oil Percentage (%)
CONTROL		24.7
CONTROL		28.0
CONTROL		31.8
CONTROL		32.4
NAPIN LCAT1	PCGN9962	28.5
NAPIN LCAT1	PCGN9962	28.9
NAPIN LCAT1	PCGN9962	29.6
NAPIN LCAT1	PCGN9962	30.1
NAPIN LCAT1	PCGN9962	30.1
NAPIN LCAT1	PCGN9962	30.1
NAPIN LCAT1	PCGN9962	30.8
NAPIN LCAT1	PCGN9962	31.0
NAPIN LCAT1	pCGN9962	32.1
NAPIN LCAT1	pCGN9962	34.2
NAPIN LCAT3	pCGN9968	26.8
NAPIN LCAT3	pCGN9968	27.4
NAPIN LCAT3	pCGN9968	29.0
NAPIN LCAT3	pCGN9968	29.0
NAPIN LCAT3	pCGN9968	32.6
NAPIN LCAT2	pCGN9983	26.5
NAPIN LCAT2	pCGN9983	34.7
NAPIN LCAT2	pCGN9983	34.8
NAPIN LCAT2	pCGN9983	35.7
NAPIN LCAT2	pCGN9983	35.8
NAPIN LCAT2	pCGN9983	36.3
NAPIN LCAT2	pCGN9983	36.7
NAPIN LCAT2	pCGN9983	37.0
NAPIN LCAT2	pCGN9983	37.2
NAPIN LCAT2	pCGN9983	37.3
NAPIN LCAT2	pCGN9983	37.3
NAPIN LCAT2	pCGN9983	37.4
NAPIN LCAT2	pCGN9983	37.8
NAPIN LCAT2	pCGN9983	38.0
NAPIN LCAT2	pCGN9983	38.0
35S LCAT2	pCGN9981	27.3
35S LCAT2	pCGN9981	28.1
35S LCAT2	pCGN9981	28.2
35S LCAT2	pCGN9981	28.6
35S LCAT2	pCGN9981	29.8
35S LCAT2	pCGN9981	30.3
35S LCAT2	pCGN9981	32.4
35S LCAT2	pCGN9981	32.5
35S LCAT2	pCGN9981	33.6
35S LCAT2	pCGN9981	34.1
35S LCAT2	pCGN9981	35.5
35S LCAT2	pCGN9981	36.4

35S LCAT2	pCGN9981	37.1
35S LCAT2	pCGN9981	38.3
35S LCAT2	pCGN9981	38.5
35S LCAT2	pCGN9981	39.1

In light of the detailed description of the invention and the examples presented above, it can be appreciated that the several aspects of the invention are achieved.

5 It is to be understood that the present invention has been described in detail by way of illustration and example in order to acquaint others skilled in the art with the invention, its principles, and its practical application. Particular formulations and processes of the present invention are not limited to the descriptions of the specific embodiments presented, but rather the descriptions and examples should be viewed in terms of the claims that follow and their  
10 equivalents. While some of the examples and descriptions above include some conclusions about the way the invention may function, the inventors do not intend to be bound by those conclusions and functions, but put them forth only as possible explanations.

It is to be further understood that the specific embodiments of the present invention as set forth are not intended as being exhaustive or limiting of the invention, and that many  
15 alternatives, modifications, and variations will be apparent to those of ordinary skill in the art in light of the foregoing examples and detailed description. Accordingly, this invention is intended to embrace all such alternatives, modifications, and variations that fall within the spirit and scope of the following claims.

00651651 083000

What is claimed is:

1. An isolated nucleic acid sequence comprising a polynucleotide encoding a plant lecithin:cholesterol acyltransferase-like polypeptide or fragment thereof.
2. The isolated nucleic acid sequence of claim 1, wherein said plant lecithin:cholesterol acyltransferase-like polypeptide is selected from the group consisting of *Arabidopsis*, soybean and corn.
3. An isolated nucleic acid sequence comprising a polynucleotide encoding a plant acyl CoA:cholesterol acyltransferase-like polypeptide.
4. The isolated nucleic acid sequence of claim 3, wherein said polynucleotide is SEQ ID NO: 42 or degenerate variants thereof.
5. The isolated nucleic acid sequence of claim 1, wherein said polynucleotide is selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 43, 44, 45, 46, 47, 48, 49, 50, 51, 73 and 75 or degenerate variants thereof.
6. An isolated nucleic acid sequence consisting essentially of SEQ ID NO: 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51 73 or 75.
7. An isolated nucleic acid sequence consisting of SEQ ID NO: 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51 73 or 75.
8. An isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of:
  - a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO 3 or SEQ ID NO 3 with at least one conservative amino acid substitution;
  - b) SEQ ID NO: 2;
  - c) an isolated polynucleotide that has at least 70% sequence identity to SEQ ID NO: 2;

- 10 d) an isolated polynucleotide that has at least 80% sequence identity to SEQ ID NO: 2;
- e) an isolated polynucleotide that has at least 90% sequence identity to SEQ ID NO: 2;
- f) an isolated polynucleotide that has at least 95% sequence identity to SEQ ID NO: 2;
- 15 g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 2;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- 20 i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 2 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
9. An isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3', where X is hydrogen, Y is hydrogen or a metal, R<sub>1</sub> and R<sub>3</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is selected from the group consisting of:
- 5 a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 3 or SEQ ID NO: 3 with at least one conservative amino acid substitution;
- b) SEQ ID NO: 2;
- c) an isolated polynucleotide that has at least 70% sequence identity to SEQ ID NO: 2;
- 10 d) an isolated polynucleotide that has at least 80% sequence identity to SEQ ID NO: 2;
- e) an isolated polynucleotide that has at least 90% sequence identity to SEQ ID NO: 2;
- f) an isolated polynucleotide that has at least 95% sequence identity to SEQ ID NO: 2;
- 15 g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 2;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and

09651651.083000



1990-1991		1991-1992		1992-1993		1993-1994		1994-1995		1995-1996		1996-1997		1997-1998		1998-1999		1999-2000		2000-2001		2001-2002		2002-2003		2003-2004		2004-2005		2005-2006		2006-2007		2007-2008		2008-2009		2009-2010		2010-2011		2011-2012		2012-2013		2013-2014		2014-2015		2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023		2023-2024		2024-2025		2025-2026		2026-2027		2027-2028		2028-2029		2029-2030		2030-2031		2031-2032		2032-2033		2033-2034		2034-2035		2035-2036		2036-2037		2037-2038		2038-2039		2039-2040		2040-2041		2041-2042		2042-2043		2043-2044		2044-2045		2045-2046		2046-2047		2047-2048		2048-2049		2049-2050		2050-2051		2051-2052		2052-2053		2053-2054		2054-2055		2055-2056		2056-2057		2057-2058		2058-2059		2059-2060		2060-2061		2061-2062		2062-2063		2063-2064		2064-2065		2065-2066		2066-2067		2067-2068		2068-2069		2069-2070		2070-2071		2071-2072		2072-2073		2073-2074		2074-2075		2075-2076		2076-2077		2077-2078		2078-2079		2079-2080		2080-2081		2081-2082		2082-2083		2083-2084		2084-2085		2085-2086		2086-2087		2087-2088		2088-2089		2089-2090		2090-2091		2091-2092		2092-2093		2093-2094		2094-2095		2095-2096		2096-2097		2097-2098		2098-2099		2099-2100		2100-2101		2101-2102		2102-2103		2103-2104		2104-2105		2105-2106		2106-2107		2107-2108		2108-2109		2109-2110		2110-2111		2111-2112		2112-2113		2113-2114		2114-2115		2115-2116		2116-2117		2117-2118		2118-2119		2119-2120		2120-2121		2121-2122		2122-2123		2123-2124		2124-2125		2125-2126		2126-2127		2127-2128		2128-2129		2129-2130		2130-2131		2131-2132		2132-2133		2133-2134		2134-2135		2135-2136		2136-2137		2137-2138		2138-2139		2139-2140		2140-2141		2141-2142		2142-2143		2143-2144		2144-2145		2145-2146		2146-2147		2147-2148		2148-2149		2149-2150		2150-2151		2151-2152		2152-2153		2153-2154		2154-2155		2155-2156		2156-2157		2157-2158		2158-2159		2159-2160		2160-2161		2161-2162		2162-2163		2163-2164		2164-2165		2165-2166		2166-2167		2167-2168		2168-2169		2169-2170		2170-2171		2171-2172		2172-2173		2173-2174		2174-2175		2175-2176		2176-2177		2177-2178		2178-2179		2179-2180		2180-2181		2181-2182		2182-2183		2183-2184		2184-2185		2185-2186		2186-2187		2187-2188		2188-2189		2189-2190		2190-2191		2191-2192		2192-2193		2193-2194		2194-2195		2195-2196		2196-2197		2197-2198		2198-2199		2199-2200		2200-2201		2201-2202		2202-2203		2203-2204		2204-2205		2205-2206		2206-2207		2207-2208		2208-2209		2209-2210		2210-2211		2211-2212		2212-2213		2213-2214		2214-2215		2215-2216		2216-2217	
-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--

- 5
- a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 5 or SEQ ID NO: 5 with at least one conservative amino acid substitution;
- b) SEQ ID NO: 4;
- c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 4;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 4;
- 10 e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 4;
- f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 4;
- 15 g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 4;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 4 and encodes a plant lecithin:cholesterol acyltransferase-like
- 20 polypeptide.

- 5 a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 5 or SEQ ID NO: 5 with at least one conservative amino acid substitution;
- b) SEQ ID NO: 4;

- 10 c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 4;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 4;
- e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 4;
- 15 f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 4;
- g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 4;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- 20 i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 4 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.

12. An isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of:

- 5 a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 7 or SEQ ID NO: 7 with at least one conservative amino acid substitution;
- b) SEQ ID NO: 6;
- c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 6;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 6;
- 10 e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 6;
- f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 6;
- g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 6;
- 15 h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and

- 20 i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 6 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
13. An isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula  $5' X-(R_1)_n-(R_2)_n-(R_3)_n-Y 3'$ , where X is hydrogen, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any nucleic acid, n is an integer between 0-3000, and  $R_2$  is selected from the group consisting of:
- 5 a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 7 or SEQ ID NO: 7 with at least one conservative amino acid substitution;
- b) SEQ ID NO: 6;
- 10 c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 6;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 6;
- e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 6;
- 15 f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 6;
- g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 6;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- 20 i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 6 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
14. An isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of
- a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 9 or SEQ ID NO: 9 with at least one conservative amino acid substitution;
- 5 b) SEQ ID NO 8;

- c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 8;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 8;
- 10 e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 8;
- f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 8;
- 15 g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 8;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 8 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
15. An isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3', where X is hydrogen, Y is hydrogen or a metal, R<sub>1</sub> and R<sub>3</sub> are any nucleic acid, n is an integer between 0-3000, and R<sub>2</sub> is selected from the group consisting of:
- 5 a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 9 or SEQ ID NO: 9 with at least one conservative amino acid substitution;
- b) SEQ ID NO: 8;
- 10 c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 8;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 8;
- e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 8;
- 15 f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 8;
- g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 8;

- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- 20 i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 8 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
16. An isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of:
- a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 74 or SEQ ID NO: 74 with at least one conservative amino acid substitution;
- 5 b) SEQ ID NO: 73;
- c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 73;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 73;
- 10 e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 73;
- f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 73;
- 15 g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 73;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- 20 i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 73 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
17. An isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula  $5' X-(R_1)_n-(R_2)_n-(R_3)_n-Y 3'$ , where X is hydrogen, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any nucleic acid, n is an integer between 0-3000, and  $R_2$  is selected from the group consisting of:
- 5 a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 74 or SEQ ID NO: 74 with at least one conservative amino acid substitution;

- 09651651.083000
- b) SEQ ID NO: 73;
- c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 73;
- 10 d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 73;
- e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 73;
- f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 73;
- 15 g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 73;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- 20 i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 73 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
18. A isolated nucleic acid sequence comprising a polynucleotide selected from the group consisting of:
- a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 76 or SEQ ID NO: 76 with at least one conservative amino acid substitution;
- 5 b) SEQ ID NO: 75;
- c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 75;
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 75;
- 10 e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 75;
- f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 75;
- 15 g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 75;

- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 75 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
- 20
19. An isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula  $5' X-(R_1)_n-(R_2)_n-(R_3)_n-Y 3'$ , where X is hydrogen, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any nucleic acid, n is an integer between 0-3000, and  $R_2$  is selected from the group consisting of:
- 5
- a) an isolated polynucleotide encoding a polypeptide of SEQ ID NO: 76 or SEQ ID NO: 76 with at least one conservative amino acid substitution;
- b) SEQ ID NO: 75;
- c) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 75;
- 10
- d) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 75;
- e) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 75;
- f) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 75;
- 15
- g) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 75;
- h) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), (f) or (g); and
- 20
- i) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 75 and encodes a plant lecithin:cholesterol acyltransferase-like polypeptide.
20. An isolated nucleic acid sequence comprising\* a polynucleotide selected from the group consisting of:
- a) SEQ ID NO: 42 or a degenerate variant thereof;

b) an isolated polynucleotide having at least 70% sequence identity with SEQ ID NO: 42;

c) an isolated polynucleotide having at least 80% sequence identity with SEQ ID NO: 42;

d) an isolated polynucleotide having at least 90% sequence identity with SEQ ID NO: 42;

e) an isolated polynucleotide having at least 95% sequence identity with SEQ ID NO: 42;

f) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 42;

g) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), or (f); and

h) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 42 and encodes an acyl CoA:cholesterol acyltransferase-like polypeptide.

21. An isolated nucleic acid sequence consisting essentially of a polynucleotide of the formula 5' X-(R<sub>1</sub>)<sub>n</sub>-(R<sub>2</sub>)<sub>n</sub>-(R<sub>3</sub>)<sub>n</sub>-Y 3', where X is hydrogen, Y is hydrogen or a metal R<sub>1</sub> and R<sub>3</sub> are any nucleic acid, n is an integer between 0 and 3000, and R<sub>2</sub> is selected from the group consisting of:

a) SEQ ID NO: 42 or degenerate variants thereof;

b) an isolated polynucleotide having at least 70% sequence identity to SEQ ID NO: 42;

c) an isolated polynucleotide having at least 80% sequence identity to SEQ ID NO: 42;

d) an isolated polynucleotide having at least 90% sequence identity to SEQ ID NO: 42;

e) an isolated polynucleotide having at least 95% sequence identity to SEQ ID NO: 42;

f) an isolated polynucleotide of at least 10 nucleic acids that hybridizes under stringent conditions to SEQ ID NO: 42;

g) an isolated polynucleotide complementary to a polynucleotide of (a), (b), (c), (d), (e), or (f); and



20

- h) an isolated polynucleotide that hybridizes under stringent conditions to SEQ ID NO: 42 and encodes an acyl CoA:cholesterol acyltransferase-like polypeptide.
22. A recombinant nucleic acid construct comprising a regulatory sequence operably linked to polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide or a fragment thereof.
23. The recombinant nucleic acid construct of claim 22, wherein said lecithin:cholesterol acyltransferase-like polypeptide is a plant lecithin:cholesterol acyltransferase-like polypeptide.
24. A recombinant nucleic acid construct comprising a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide.
25. The recombinant nucleic acid construct of claim 24, wherein said acyl CoA:cholesterol acyltransferase-like polypeptide is a plant acyl CoA:cholesterol acyltransferase-like polypeptide.
26. The recombinant construct of claim 22, wherein said regulatory sequence comprises a heterologous regulatory sequence.
27. The recombinant construct of claim 24, wherein said regulatory sequence comprises a heterologous regulatory sequence.
28. The recombinant construct of claim 22, wherein said regulatory sequence is functional in a plant cell.
29. The recombinant construct of claim 24, wherein said regulatory sequence is functional in a plant cell.
30. The recombinant construct of claim 22, further comprising a termination sequence.

31. The recombinant construct of claim 24 further comprising a termination sequence.
32. The recombinant construct of claim 22 wherein said polynucleotide is selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 43, 44, 45, 46, 47, 48, 49, 50, 51, 73 and 75.
33. The recombinant construct of claim 24, wherein said polynucleotide is selected from the group consisting of SEQ ID NO: 33 and 42.
34. The recombinant construct of claim 22, wherein said regulatory sequence comprises a constitutive promoter.
35. The recombinant construct of claim 24, wherein said regulatory sequence comprises a constitutive promoter.
36. The recombinant construct of claim 22, wherein said regulatory sequence comprises an inducible promoter.
37. The recombinant construct of claim 24, wherein said regulatory sequence comprises an inducible promoter.
38. The recombinant construct of claim 22, wherein said regulatory sequence is selected from the group consisting of a tissue specific promoter, a developmentally regulated promoter, an organelle specific promoter, and a seed specific promoter.
39. The recombinant construct of claim 24, wherein said regulatory sequence is selected from the group consisting of a tissue specific promoter, a developmentally regulated promoter, an organelle specific promoter, and a seed specific promoter.
40. A host cell containing the recombinant construct of claim 22 or 24.

41. The host cell of claim 40, wherein said host cell is selected from the group consisting of plant cells, animal cells, insect cells, yeast, bacteria, bacteriophage and viruses.
42. The host cell of claim 40, wherein said host cell is a plant cell.
43. The host cell of claim 40, wherein said host cell expresses a lecithin:cholesterol acyltransferase-like polypeptide or an acyl CoA:cholesterol acyltransferase-like polypeptide.
44. The host cell of claim 43, wherein said cholesterol acyltransferase-like polypeptide is a plant acyltransferase-like polypeptide.
45. A plant comprising at least one host cell of claim 40.
46. The progeny of a plant of claim 45.
47. A seed from the plant of claim 45.
48. A plant comprising the recombinant construct of claim 22 or 24.
49. The progeny of a plant of claim 48.
50. A seed from the plant of claim 48.
51. A purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 3 with at least one conservative amino acid substitution, SEQ ID NO: 5, SEQ ID NO: 5 with at least one conservative amino acid substitution, SEQ ID NO: 7, SEQ ID NO: 7 with at least one conservative amino acid substitution, SEQ ID NO: 9, SEQ ID NO: 9 with at least one conservative amino acid substitution, SEQ ID NO: 74, SEQ ID NO: 74 with at least one conservative amino acid substitution, SEQ ID NO: 76 and SEQ ID NO: 76 with at least one conservative amino acid substitution.

52. A purified immunogenic polypeptide comprising at least 10 consecutive amino acids from an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 3 with at least one conservative amino acid substitution, SEQ ID NO: 5, SEQ ID NO: 5 with at least one conservative amino acid substitution, SEQ ID NO: 7, SEQ ID NO: 7 with at least one conservative amino acid substitution, SEQ ID NO: 9, SEQ ID NO: 9 with at least one conservative amino acid substitution, SEQ ID NO: 74, SEQ ID NO: 74 with at least one conservative amino acid substitution, SEQ ID NO: 76 and SEQ ID NO: 76 with at least one conservative amino acid substitution.
53. An antibody which specifically binds to an immunogenic polypeptide of claim 52.
54. A method for producing a lecithin:cholesterol acyltransferase-like polypeptide or an acyl CoA:cholesterol acyltransferase-like polypeptide comprising culturing a host cell of claim 40 under conditions permitting expression of said lecithin:cholesterol acyltransferase-like polypeptide or acyl CoA:cholesterol acyltransferase-like polypeptide.
55. The method of claim 54, further comprising isolating the cholesterol acyltransferase-like polypeptide from the host cell or from the medium in which the host cell is cultured.
56. A method for modifying the sterol content of a host cell, comprising transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide and culturing said host cell under conditions wherein said host cell expresses a lecithin:cholesterol acyltransferase-like polypeptide such that said host cell has a modified sterol composition as compared to host cells without the recombinant construct.
57. The method of claim 56, wherein said lecithin:cholesterol acyltransferase-like polypeptide is a plant lecithin:cholesterol acyltransferase-like polypeptide.

58. A method for modifying the sterol content of a host cell, comprising transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide and culturing said host cell under conditions wherein said host cell expresses an acyl CoA:cholesterol acyltransferase-like polypeptide such that said host cell has a modified sterol composition as compared to host cells without the recombinant construct.
59. The method of claim 58, wherein said acyl CoA:cholesterol acyltransferase-like polypeptide is a plant acyl CoA:cholesterol acyltransferase-like polypeptide.
60. The method of claim 56, wherein said modified sterol composition is an increase in sterol esters.
61. The method of claim 58, wherein said modified sterol composition is an increase in sterol esters.
62. The method of claim 56, wherein said polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide is selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 11, 73 and 75.
63. The method of claim 58 wherein said polynucleotide encoding a acyl CoA:cholesterol acyltransferase-like polypeptide is SEQ ID NO 33 OR 42.
64. The method of claim 56, wherein said regulatory sequence comprises a constitutive promoter.
65. The method of claim 58, wherein said regulatory sequence comprises a constitutive promoter.
66. The method of claim 56, wherein said regulatory sequence is an inducible promoter.
67. The method of claim 58, wherein said regulatory sequence is an inducible promoter.

68. The method of claim 56, wherein said regulatory sequence is a tissue specific promoter.
69. The method of claim 58, wherein said regulatory sequence is a tissue specific promoter.
70. The method of claim 56, wherein said regulatory sequence is a seed specific promoter.
71. The method of claim 58, wherein said regulatory sequence is a seed specific promoter.
72. The method of claim 56, wherein said polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide is in the antisense orientation.
73. The method of claim 58, wherein said polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide is in the antisense orientation.
74. The method of claim 72, wherein said modified sterol composition is a decrease in sterol esters.
75. The method of claim 73, wherein said modified sterol composition is a decrease in sterol esters.
76. A plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in modified sterol composition of said plant as compared to the same plant without said recombinant construct.
77. The plant of claim 76, wherein said lecithin:cholesterol acyltransferase-like polypeptide is a plant lecithin:cholesterol acyltransferase-like polypeptide.

78. The plant of claim 76, wherein said polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide is selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 11, 73 and 75.
79. A plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in modified sterol composition of said plant as compared to the same plant without said recombinant construct.
80. The plant of claim 79, wherein said acyl CoA:cholesterol acyltransferase-like polypeptide is a plant acyl CoA:cholesterol acyltransferase-like polypeptide.
81. The plant of claim 79, wherein said polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide is SEQ ID NO: 33 or 42.
82. The plant of claim 76, wherein said regulatory sequence comprises a tissue specific promoter.
83. The plant of claim 79, wherein said regulatory sequence comprises a tissue specific promoter.
84. The plant of claim 76, wherein said regulatory sequence comprises a seed specific promoter.
85. The plant of claim 79, wherein said regulatory sequence comprises a seed specific promoter.
86. The plant of claim 76, wherein said modified sterol composition is an increase in sterol esters.
87. The plant of claim 79, wherein said modified sterol composition is an increase in sterol esters.

88. The plant of claim 76, wherein the polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide is in the antisense orientation.
89. The plant of claim 79, wherein the polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide is in the antisense orientation.
90. An oil obtained from the plant of claim 76 or 79.
91. A method for producing an oil with a modified sterol composition comprising, providing a plant of claim 76 or 79 and extracting the oil from said plant.
92. An oil produced by the method of claim 91.
93. A method for altering oil production by a host cell comprising, transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide and culturing said host cell under conditions wherein said host cell expresses a  
5 lecithin:cholesterol acyltransferase-like polypeptide such that said host cell has an altered oil production as compared to host cells without the recombinant construct.
94. The method of claim 93, wherein said lecithin:cholesterol acyltransferase-like polypeptide is a plant lecithin:cholesterol acyltransferase-like polypeptide.
95. A method for altering oil production by a host cell comprising, transforming a host cell with a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide and culturing said host cell under conditions wherein said host cell expresses an acyl  
5 CoA:cholesterol acyltransferase-like polypeptide such that said host cell has an altered oil production as compared to host cells without the recombinant construct.
96. The method of claim 95, wherein said acyl CoA:cholesterol acyltransferase-like polypeptide is a plant acyl CoA:cholesterol acyltransferase-like polypeptide.



97. The method of claim 93, wherein said oil production is increased.
98. The method of claim 95, wherein said oil production is increased.
99. The method of claim 93, wherein said host cell is a plant cell.
100. The method of claim 95, wherein said host cell is a plant cell.
101. The method of claim 93, wherein said polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide is selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 11, 73 and 75.
102. The method of claim 95, wherein said polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide is SEQ ID NO: 33 or 42.
103. The method of claim 93, wherein said regulatory sequence is a tissue specific promoter.
104. The method of claim 95, wherein said regulatory sequence is a tissue specific promoter.
105. The method of claim 93, wherein said regulatory sequence is a seed specific promoter.
106. The method of claim 95, wherein said regulatory sequence is a seed specific promoter.
107. A plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in an altered production of oil by said plant as compared to the same plant without said recombinant construct.

108. The plant of claim 107, wherein said lecithin:cholesterol acyltransferase-like polypeptide is a plant lecithin:cholesterol acyltransferase-like polypeptide.
109. A plant comprising a recombinant construct containing a regulatory sequence operably linked to a polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide wherein expression of said recombinant construct results in an altered production of oil by said plant as compared to the same plant without said recombinant construct.
110. The plant of claim 109, wherein said acyl CoA:cholesterol acyltransferase-like polypeptide is a plant acyl CoA:cholesterol acyltransferase-like polypeptide.
111. The plant of claim 107, wherein said oil production is increased.
112. The plant of claim 109, wherein said oil production is increased.
113. The plant of claim 107, wherein said polynucleotide encoding a lecithin:cholesterol acyltransferase-like polypeptide is selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 11, 73 and 75.
114. The plant of claim 109, wherein said polynucleotide encoding an acyl CoA:cholesterol acyltransferase-like polypeptide is SEQ ID NO: 33 or 42.
115. The plant of claim 107, wherein said regulatory sequence is a tissue specific promoter.
116. The plant of claim 109, wherein said regulatory sequence is a tissue specific promoter.
117. The plant of claim 107, wherein said regulatory sequence is a seed specific promoter.
118. The plant of claim 109, wherein said regulatory sequence is a seed specific promoter.
119. A food, food ingredient or food product comprising the oil of claim 90 or 92.

120. A food, food ingredient or food product comprising the plant of claim 107 or 109.

[illegible]

4

# ClustalW Formatted Alignments

Yeast (YNR008W) G S A K R N E R G K D F D R K R D G N G R K R W R D S R R L I F I L G A F L G V L L P F S F G A Y H  
Human LCAT  
Rat LCAT  
At LCAT1  
At LCAT2  
At LCAT3  
At LCAT4

	110	120	130	140	150
Yeast (YNR008W)	V H N S D S D L F D	N F V N F D S L K V Y L D D W K D	V L P Q G	I S S F I D D I Q A G N Y S T S S L	
Human LCAT	F P P H T T P K A E	L S N - - - H T R	P V I L V P G	C L - - - - -	
Rat LCAT	F P P H T T P K A E	L S N - - - H T R	P V I L V P G	C M - - - - -	
Ai LCAT1	M C Q A V G S N V Y	- - - - -	P L I L V P G	N G - - - - -	
Ai LCAT2	T A V E D E T E F H G D Y	- - - - -	S K L S G I I I P G F A	- - - - -	
Ai LCAT3	Q E P Y V D P N L N	- - - - -	P V L L V P G I A	- - - - -	
Ai LCAT4	E N A G E V A D R D	- - - - -	P V L L V S G I G	- - - - -	
			P V I L V P G		

FIG. 1B

Yeast (YNR008W)	160	170	180	190	200
Human LCAT	D D L S E N F A V	E A K H P	- - - V M V P G V I S T G I	E S W G V	I G D D
Rat LCAT	- - - - -	D K P D V	- - - V N W M C	- - - Y R K T E	I W L D
AtLCAT1	- - - - -	D K P N V	- - - V N W L C	- - - Y R K T E	I W L D
AtLCAT2	- - - - -	D R E Y K P S S V W C	- - - T P	- - - L D F N P L D L	L W F D
AtLCAT3	- - - - -	D C P Y	- - - N	- - - G N E R	V W V R
AtLCAT4	- - - - -	D H E	- - - S	- - - K S E I R	V W V R
		D		K . E	W . D

Yeast (YNR008W)	210	220	230	240	250
Human LCAT	E C D S	L N M F	L P L G V D	L W G S F Y M	L R T M
Rat LCAT	- - - - -	F N M F	L P L G V D	C W I D N T R V	- - - - -
AtLCAT1	- - - - -	A A V L	L S P F T R	C F S D R M M L	- - - - -
AtLCAT2	- - - - -	T T K L	L S A V N C	W F K - - C M	- - - - -
AtLCAT3	- - - - -	I F G A	D H E F R T K M	W S - - - -	- - - - -
AtLCAT4	- - - - -	I F L A	N L A F K Q S L	W S - - - -	- - - - -
		L			

Yeast (YNR008W)	260	270	280	290	300
Human LCAT	A A Q G F E S T D Y F I A G	- - - - -	Y W I W N	K V F Q N	L G V I
Rat LCAT	Y S V E Y L D S S K L A G Y	- - - - -	L H T L V Q N	L V N N	L V N N
AtLCAT1	Y S V E Y L D S S K L A G Y	- - - - -	L N T L V Q N	L V N N	L V N N
AtLCAT2	K S L L Y L D P P R L R D A T	- - - - -	S Y M E H L	V K A L E K K C	G Y V N D Q T
AtLCAT3	S A I T F L D P P G Y I T G P	- - - - -	L S T V W K E W L	K W C V E F	G I E A N
AtLCAT4	H A I D V L D P P D M I V G R	- - - - -	E S V Y Y F H	E M I V E M	G F F E E G K T
	Y A I D I L D P P S W F V K L C H L T E V Y H F H		H		
	A . . . L D P				

Yeast (YNR008W)

FIG. 2

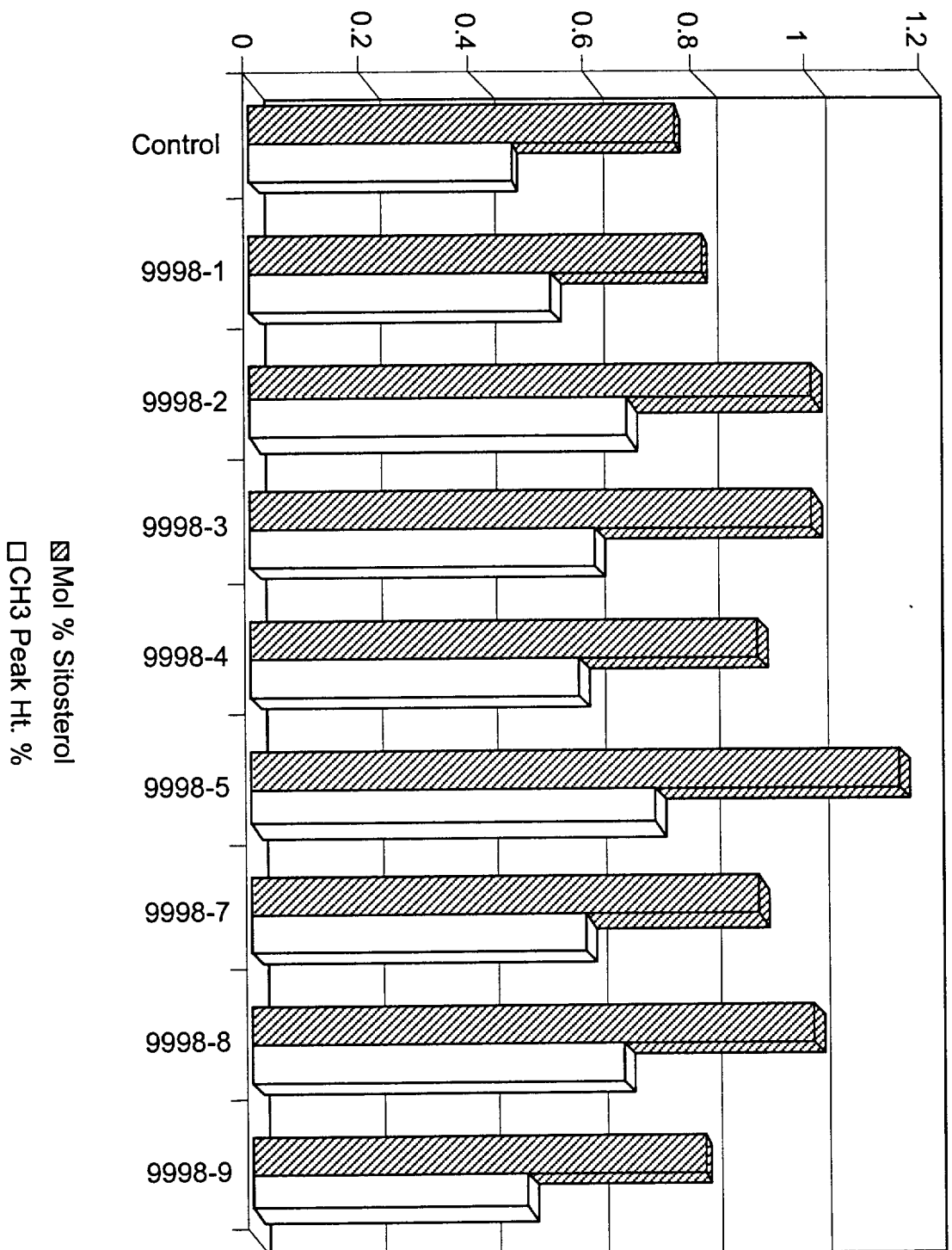
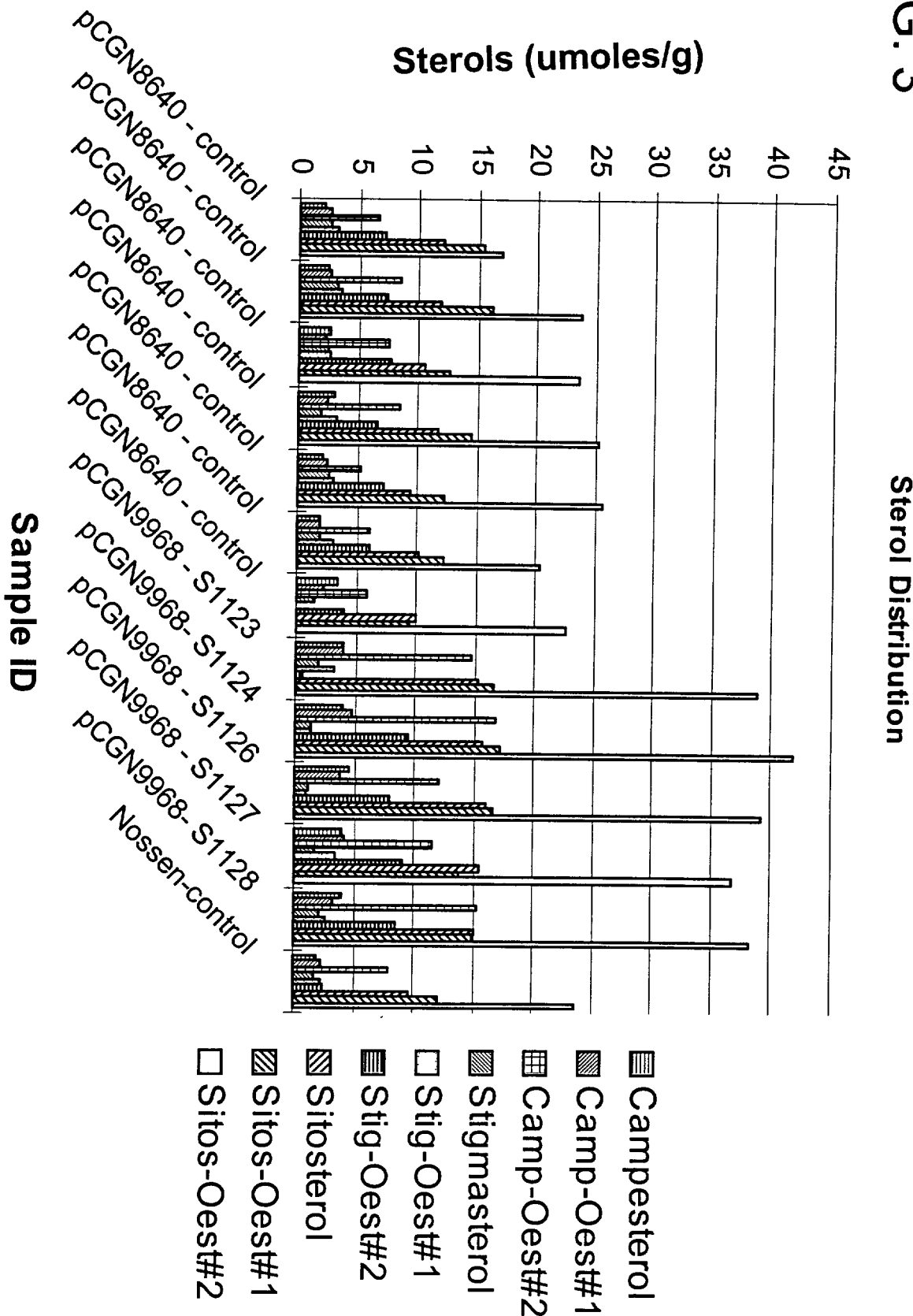
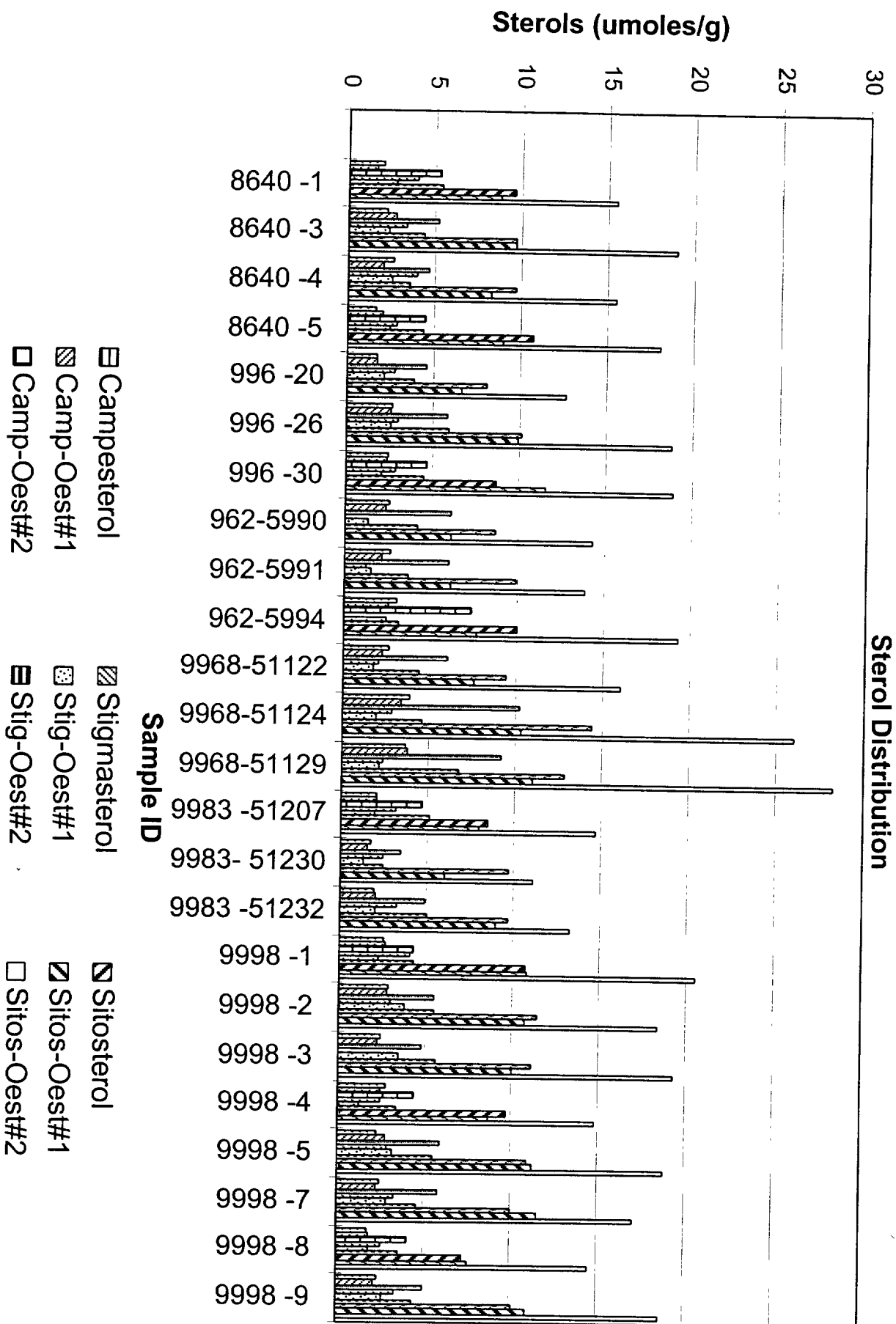




FIG. 3



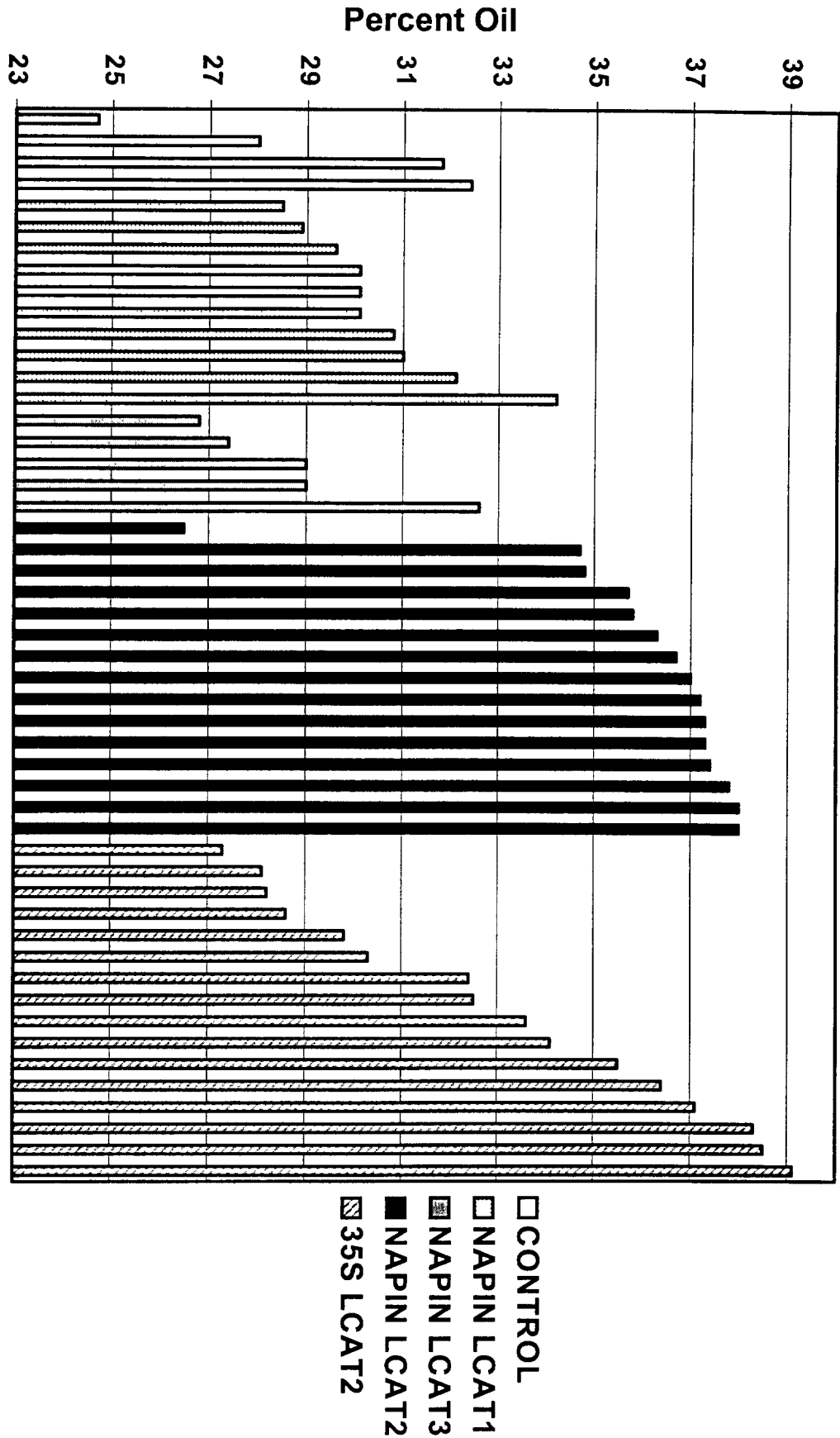
# FIG. 4



34743

FIG. 5

NIR Analysis of LCAT



**DECLARATION AND POWER OF ATTORNEY**

**REGULAR OR DESIGN APPLICATION**

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

PLANT STEROL ACYLTRANSFERASES

the specification of which:

(check one)

☐ is attached hereto

☐ was filed on \_\_\_\_\_ as Application Serial No. \_\_\_\_\_, and was amended on \_\_\_\_\_.

☐ was described and claimed in PCT International Application No. \_\_\_\_\_, filed on \_\_\_\_\_ and as amended under PCT Article 19 on \_\_\_\_\_, if any.

**ACKNOWLEDGEMENT OF REVIEW OF PAPERS AND DUTY OF CANDOR**

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations §1.56.

[illegible]

Priority Claimed

(Number)	(Country)	(Day/Month/Year Filed)
----------	-----------	------------------------

ANY FOREIGN APPLICATION(S), ON THE SAME SUBJECT MATTER WHICH HAS  
A FILING DATE EARLIER THAN THE EARLIEST APPLICATION FROM WHICH  
PRIORITY IS CLAIMED

CLAIM FOR BENEFIT OF PROVISIONAL APPLICATION(S)

(Application Number)	(Filing Date)
----------------------	---------------

**CLAIM FOR BENEFIT OF EARLIER U.S. APPLICATION(S)  
UNDER 35 U.S.C. 120**

(complete this part only if this is a divisional,  
continuation or CIP application)

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s), or §365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

_____ (Serial No.)	_____ (Filing Date)	_____ (Status)
_____ (Serial No.)	_____ (Filing Date)	_____ (Status)

**POWER OF ATTORNEY**

I hereby appoint the following attorneys to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Irving Powers (15,700), Donald G. Leavitt (17,626), John K. Roedel, Jr. (25,914), Michael E. Godar (28,416), Edward J. Hejlek (31,525), William E. Lahey (26,757), Richard G. Heywood (18,224), Frank R. Agovino (27,416), Kurt F. James (33,716), G. Harley Blosser (33,650), Paul I. J. Fleischut (35,513), Vincent M. Keil (36,838), Robert M. Evans, Jr. (36,794), Robert M. Bain (36,736), Joseph A. Schaper (30,493), Kathleen M. Petrillo (35,076), David E. Crawford, Jr. (38,118), Paul A. Maddock (37,877), Richard L. Bridge (40,529), Christopher M. Goff (41,785), James E. Butler (40,931), Derick E. Allen (43,468), Matthew L. Cutler (43,574), Michael G. Munsell (43,820), Karen Y. Hui (44,785), Anthony R. Kinney (44,834), Brian P. Klein (44,837), Sarah J. Chickos (46,157), Donald W. Tuegel (45,424), Steven M. Ritchey (46,321), Michael J. Thomas (39,857), Kathryn J. Doty (40,593), and Laura R. Polcyn (P-47,000), all of the law firm of SENNIGER, POWERS, LEAVITT & ROEDEL, One Metropolitan Square, 16th Floor, St. Louis, Missouri 63102.

Send Correspondence To: Direct Telephone Calls To:

Customer Number: 000321 James E. Butler, Ph.D.  
(314) 231-5400

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor Michael Lassner

Inventor's signature \_\_\_\_\_ Date \_\_\_\_\_

Residence Redwood City, CA Citizenship U.S.A.

Post Office address 515 Galveston Drive

Redwood City, California 94063

Full name of second joint inventor Alison Van Eenennaam

Second inventor's signature \_\_\_\_\_ Date \_\_\_\_\_

Residence Davis, California Citizenship Australia

Post Office address 856 Burr Street

Davis, California 95616

a) $\text{H}_2\text{O}$ and $\text{H}_2\text{O}_2$ concentrations		b) $\text{H}_2\text{O}$ and $\text{H}_2\text{O}_2$ concentrations	
Time (h)	Concentration (M)	Time (h)	Concentration (M)
0	0.00	0	0.00
1	0.01	1	0.01
2	0.02	2	0.02
3	0.03	3	0.03
4	0.04	4	0.04
5	0.05	5	0.05
6	0.06	6	0.06
7	0.07	7	0.07
8	0.08	8	0.08
9	0.09	9	0.09
10	0.10	10	0.10
11	0.11	11	0.11
12	0.12	12	0.12
13	0.13	13	0.13
14	0.14	14	0.14
15	0.15	15	0.15
16	0.16	16	0.16
17	0.17	17	0.17
18	0.18	18	0.18
19	0.19	19	0.19
20	0.20	20	0.20
21	0.21	21	0.21
22	0.22	22	0.22
23	0.23	23	0.23
24	0.24	24	0.24
25	0.25	25	0.25
26	0.26	26	0.26
27	0.27	27	0.27
28	0.28	28	0.28
29	0.29	29	0.29
30	0.30	30	0.30
31	0.31	31	0.31
32	0.32	32	0.32
33	0.33	33	0.33
34	0.34	34	0.34
35	0.35	35	0.35
36	0.36	36	0.36
37	0.37	37	0.37
38	0.38	38	0.38
39	0.39	39	0.39
40	0.40	40	0.40
41	0.41	41	0.41
42	0.42	42	0.42
43	0.43	43	0.43
44	0.44	44	0.44
45	0.45	45	0.45
46	0.46	46	0.46
47	0.47	47	0.47
48	0.48	48	0.48
49	0.49	49	0.49
50	0.50	50	0.50
51	0.51	51	0.51
52	0.52	52	0.52
53	0.53	53	0.53
54	0.54	54	0.54
55	0.55	55	0.55
56	0.56	56	0.56
57	0.57	57	0.57
58	0.58	58	0.58
59	0.59	59	0.59
60	0.60	60	0.60
61	0.61	61	0.61
62	0.62	62	0.62
63	0.63	63	0.63
64	0.64	64	0.64
65	0.65	65	0.65
66	0.66	66	0.66
67	0.67	67	0.67
68	0.68	68	0.68
69	0.69	69	0.69
70	0.70	70	0.70
71	0.71	71	0.71
72	0.72	72	0.72
73	0.73	73	0.73
74	0.74	74	0.74
75	0.75	75	0.75
76	0.76	76	0.76
77	0.77	77	0.77
78	0.78	78	0.78
79	0.79	79	0.79
80	0.80	80	0.80
81	0.81	81	0.81
82	0.82	82	0.82
83	0.83	83	0.83
84	0.84	84	0.84
85	0.85	85	0.85
86	0.86	86	0.86
87	0.87	87	0.87
88	0.88	88	0.88
89	0.89	89	0.89
90	0.90	90	0.90
91	0.91	91	0.91

Leu Val Ser Asn Ala Pro Gly Val Gln Ile Arg Val Pro Gly Phe Gly  
115 120 125







[illegible]

Met	Lys	Lys	Ile	Ser	Ser	His	Tyr	Ser	Val	Val	Ile	Ala	Ile	Leu	Val
1				5					10					15	
Val	Val	Thr	Met	Thr	Ser	Met	Cys	Gln	Ala	Val	Gly	Ser	Asn	Val	Tyr
			20					25					30		
Pro	Leu	Ile	Leu	Val	Pro	Gly	Asn	Gly	Gly	Asn	Gln	Leu	Glu	Val	Arg
		35					40					45			
Leu	Asp	Arg	Glu	Tyr	Lys	Pro	Ser	Ser	Val	Trp	Cys	Ser	Ser	Trp	Leu
	50					55					60				
Tyr	Pro	Ile	His	Lys	Lys	Ser	Gly	Gly	Trp	Phe	Arg	Leu	Trp	Phe	Asp
65					70					75					80
Ala	Ala	Val	Leu	Leu	Ser	Pro	Phe	Thr	Arg	Cys	Phe	Ser	Asp	Arg	Met
				85					90					95	
Met	Leu	Tyr	Tyr	Asp	Pro	Asp	Leu	Asp	Asp	Tyr	Gln	Asn	Ala	Pro	Gly
			100					105					110		
Val	Gln	Thr	Arg	Val	Pro	His	Phe	Gly	Ser	Thr	Lys	Ser	Leu	Leu	Tyr
		115					120					125			
Leu	Asp	Pro	Arg	Leu	Arg	Asp	Ala	Thr	Ser	Tyr	Met	Glu	His	Leu	Val
	130					135					140				
Lys	Ala	Leu	Glu	Lys	Lys	Cys	Gly	Tyr	Val	Asn	Asp	Gln	Thr	Ile	Leu
145					150					155					160
Gly	Ala	Pro	Tyr	Asp	Phe	Arg	Tyr	Gly	Leu	Ala	Ala	Ser	Gly	His	Pro
				165					170					175	
Ser	Arg	Val	Ala	Ser	Gln	Phe	Leu	Gln	Asp	Leu	Lys	Gln	Leu	Val	Glu
			180					185					190		
Lys	Thr	Ser	Ser	Glu	Asn	Glu	Gly	Lys	Pro	Val	Ile	Leu	Leu	Ser	His
		195					200					205			
Ser	Leu	Gly	Gly	Leu	Phe	Val	Leu	His	Phe	Leu	Asn	Arg	Thr	Thr	Pro
	210					215					220				
Ser	Trp	Arg	Arg	Lys	Tyr	Ile	Lys	His	Phe	Val	Ala	Leu	Ala	Ala	Pro
225					230					235					240
Trp	Gly	Gly	Thr	Ile	Ser	Gln	Met	Lys	Thr	Phe	Ala	Ser	Gly	Asn	Thr
				245					250					255	















				325					330					335	
Val	Leu	Ala	Ser	Ala	Lys	Leu	Pro	Pro	Lys	Val	Lys	Phe	Tyr	Asn	Ile
			340					345					350		
Tyr	Gly	Thr	Asn	Leu	Glu	Thr	Pro	His	Ser	Val	Cys	Tyr	Gly	Asn	Glu
		355					360					365			
Lys	Met	Pro	Val	Lys	Asp	Leu	Thr	Asn	Leu	Arg	Tyr	Phe	Gln	Pro	Thr
	370					375					380				
Tyr	Ile	Cys	Val	Asp	Gly	Asp	Gly	Thr	Val	Pro	Met	Glu	Ser	Ala	Met
385					390					395					400
Ala	Asp	Gly	Leu	Glu	Ala	Val	Ala	Arg	Val	Gly	Val	Pro	Gly	Glu	His
				405					410					415	
Arg	Gly	Ile	Leu	Asn	Asp	His	Arg	Val	Phe	Arg	Met	Leu	Lys	Lys	Trp
			420					425					430		
Leu	Asn	Val	Gly	Glu	Pro	Asp	Pro	Phe	Tyr	Asn	Pro	Val	Asn	Asp	Tyr
		435					440					445			
Val	Ile	Leu	Pro	Thr	Thr	Tyr	Glu	Phe	Glu	Lys	Phe	His	Glu	Asn	Gly
	450					455					460				
Leu	Glu	Val	Ala	Ser	Val	Lys	Glu	Ser	Trp	Asp	Ile	Ile	Ser	Asp	Asp
465					470					475					480
Asn	Asn	Ile	Gly	Thr	Thr	Gly	Ser	Thr	Val	Asn	Ser	Ile	Ser	Val	Ser
				485					490					495	
Gln	Pro	Gly	Asp	Asp	Gln	Asn	Pro	Gln	Ala	Glu	Ala	Arg	Ala	Thr	Leu
			500					505					510		
Thr	Val	Gln	Pro	Gln	Ser	Asp	Gly	Arg	Gln	His	Val	Glu	Leu	Asn	Ala
		515					520					525			
Val	Ser	Val	Ser	Val	Asp	Ala									
	530					535									

<210> 8

<212> DNA

<400> 8

Variable	Mean	SD	Min	Max
Age	35.2	10.5	18	65
Gender	1.2	0.4	1	2
Marital status	1.5	0.5	1	3
Education	12.5	1.5	9	16
Income	3.5	1.5	1	7
Occupation	2.5	1.5	1	5
Health status	1.5	0.5	1	2
Stress level	3.5	1.5	1	7
Life satisfaction	4.5	1.5	1	7
Resilience	3.5	1.5	1	7
Optimism	4.5	1.5	1	7
Gratitude	4.5	1.5	1	7
Self-compassion	4.5	1.5	1	7
Emotional regulation	4.5	1.5	1	7
Prosocial behavior	4.5	1.5	1	7
Life meaning	4.5	1.5	1	7
Existential well-being	4.5	1.5	1	7
Psychological well-being	4.5	1.5	1	7
Overall well-being	4.5	1.5	1	7

<211> 447

<213> Arabidopsis thaliana

Met Gly Trp Ile Pro Cys Pro Cys Trp Gly Thr Asn Asp Asp Glu Asn  
1 5 10 15

Gly Gly Ser Ile Leu His Ser Lys Lys Lys Asn Ser Lys Ser Glu Ile  
35 40 45

Leu Trp Ser Leu Tyr Asn Pro Lys Thr Gly Tyr Thr Glu Pro Leu Asp  
65 70 75 80

Asp Asn Ile Glu Val Leu Val Pro Asp Asp Asp His Gly Leu Tyr Ala  
85 90 95







2000-2001		2001-2002		2002-2003		2003-2004		2004-2005		2005-2006		2006-2007		2007-2008		2008-2009		2009-2010		2010-2011		2011-2012		2012-2013		2013-2014		2014-2015		2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023		2023-2024		2024-2025		2025-2026		2026-2027		2027-2028		2028-2029		2029-2030		2030-2031		2031-2032		2032-2033		2033-2034		2034-2035		2035-2036		2036-2037		2037-2038		2038-2039		2039-2040		2040-2041		2041-2042		2042-2043		2043-2044		2044-2045		2045-2046		2046-2047		2047-2048		2048-2049		2049-2050		2050-2051		2051-2052		2052-2053		2053-2054		2054-2055		2055-2056		2056-2057		2057-2058		2058-2059		2059-2060		2060-2061		2061-2062		2062-2063		2063-2064		2064-2065		2065-2066		2066-2067		2067-2068		2068-2069		2069-2070		2070-2071		2071-2072		2072-2073		2073-2074		2074-2075		2075-2076		2076-2077		2077-2078		2078-2079		2079-2080		2080-2081		2081-2082		2082-2083		2083-2084		2084-2085		2085-2086		2086-2087		2087-2088		2088-2089		2089-2090		2090-2091		2091-2092		2092-2093		2093-2094		2094-2095		2095-2096		2096-2097		2097-2098		2098-2099		2099-2100		2100-2101		2101-2102		2102-2103		2103-2104		2104-2105		2105-2106		2106-2107		2107-2108		2108-2109		2109-2110		2110-2111		2111-2112		2112-2113		2113-2114		2114-2115		2115-2116		2116-2117		2117-2118		2118-2119		2119-2120		2120-2121		2121-2122		2122-2123		2123-2124		2124-2125		2125-2126		2126-2127		2127-2128		2128-2129		2129-2130		2130-2131		2131-2132		2132-2133		2133-2134		2134-2135		2135-2136		2136-2137		2137-2138		2138-2139		2139-2140		2140-2141		2141-2142		2142-2143		2143-2144		2144-2145		2145-2146		2146-2147		2147-2148		2148-2149		2149-2150		2150-2151		2151-2152		2152-2153		2153-2154		2154-2155		2155-2156		2156-2157		2157-2158		2158-2159		2159-2160		2160-2161		2161-2162		2162-2163		2163-2164		2164-2165		2165-2166		2166-2167		2167-2168		2168-2169		2169-2170		2170-2171		2171-2172		2172-2173		2173-2174		2174-2175		2175-2176		2176-2177		2177-2178		2178-2179		2179-2180		2180-2181		2181-2182		2182-2183		2183-2184		2184-2185		2185-2186		2186-2187		2187-2188		2188-2189		2189-2190		2190-2191		2191-2192		2192-2193		2193-2194		2194-2195		2195-2196		2196-2197		2197-2198		2198-2199		2199-2200		2200-2201		2201-2202		2202-2203		2203-2204		2204-2205		2205-2206		2206-2207		2207-2208		2208-2209		2209-2210		2210-2211		2211-2212		2212-2213		2213-2214		2214-2215		2215-2216		2216-2217		2217-2218		2218-2219		2219-2220		2220-2221		2221-2222		2222-2223		2223-2224		2224-2225		2225-2226		2226-2227	
-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--

```
<220>  
<221> unsure  
<222> (241)  
<223> n=unknown
```

```
<400> 12
ccaagaactc gatgattact tcaacactcc tggggttgng acccgggtcc ctcactttgg 60
ttccaccaac tctctttctc catctcaatc ctctgtctca gcatatcacc ggatacatgg 120
```







Table 1. Demographic characteristics of the study population	
Age (years)	65.0 ± 1.5
Gender (male/female)	10/10
Education (years)	12.0 ± 1.0
Occupation (white/blue)	10/10
Marital status (married/divorced/widowed)	10/10/0
Smoking status (smoker/nonsmoker)	10/10
Alcohol consumption (yes/no)	10/10
Comorbidities (hypertension/diabetes/cholesterol)	10/10/10
Medication (antihypertensive/antidiabetic/anticholesterol)	10/10/10
Physical activity (yes/no)	10/10
Stress level (low/high)	10/10
Sleep quality (good/poor)	10/10
Depression score (0-10)	5.0 ± 1.0
Overall health status (good/fair/poor)	10/10/0

<212> DNA

<213> Glycine max

<400> 18

gatgaaacta aaccgtgggc gactaagctt gtttactcgg ttgatttatg gcaagatcaa 60  
gttcgttgct tcatagaaga ggtcattggg gaaccagtct atcttggtgg caactcacta 120  
ggaggattgg ttgcattgta ttttgcgga aacaaccctc atttagtgaa aggtgtcgca 180  
ttgcttaagc aacacctttt tgggggtttc tgccaaatcc cataaaaagt ccaagactag 240  
cgaaaatatt tccatgggcc gga 263

<210> 19

<211> 311

<212> DNA

<213> Zea mays

<220>

<221> unsure

<222> (1)..(311)

<223> n=unknown

<400> 19

cggacgtcgg ncatgttcgg agccccctac gacttccgct acgcgccgcc gtcccccgcc 60  
cagacgtccg aggtgtactc ccgctacttc aaggagctga tggagctggg cgaggccgcg 120  
agcgagagga cccggaagaa ggccgtcatc ctccggccaca gcttcggcgg catgggtcgcg 180  
ctcgagttcg tccggaacac tccgccggcg tggcgccgcg agcacatcga gcgcctcgtc 240  
ctggtcgcgc cgacgtccc cggcggttcc ctggagccgg tgcgcaactt cgcgtccggg 300  
acggacatcc t 311

<210> 20

<211> 1155

<212> DNA

<213> Zea mays

<400> 20

tgcacccacg cgtccggcca caagaaccct ctcaagtcag actggtgcct cggaaagctg 60  
agagccgcac tgggaagacat gggataccga gacggagaca ccatgttcgg agccccctac 120  
gacttccgct acgcgccgcc gtcccccgcc cagacgtccg aggtgtactc ccgctacttc 180  
aaggagctga tggagctggg cgaggccgca agcgagagga cccggaagaa ggccgtcatc 240  
ctccggccaca gcttcggcgg catgggtcgcg ctccggaacac tccgccggcg 300  
tggcgccgcg agcacatcga gcgcctcgtc ctgggtcgcgc cgacgtccc cggcggttcc 360  
ctggagccgg tgcgcaactt cgcgtccggg acggacatcc tmtacgtgcc agcgacgacg 420  
ccgtggcca cgcgagccat gtgragragc ttcgagagcg ccatcgtgaa cttcccgtcg 480  
ccggccgtgt tccggcgccct gcaggcgccg ctccgtgtca ccaggagcg gaactactcc 540  
gcgtccgcgc acgacatgga gcgttccctc gccgccgctc gctccggcga ggccgcggag 600  
cccttcagga gacgggcccgt ccccaagatg ggcagcttcg cggcgccgat ggtgcccag 660  
acgtacatca gcggggtcgg caacaggacg ccgctgcggc tgggtgttctg gggcgacgac 720  
ttcgacgcgg ccccgagggt ggccggcgta ggggacggag atggcaagat caatttgatc 780





<223> n=unknown

<400> 25

ggagattgtc gtgccggagg acgaccacgg cctgtttgcc atcgacattc ttgatccttc 60  
ctggtttgta gaactcgacc cacggtccg cccaccgtcc gggagattgt cgtgccggag 120  
gacgaccacg gcctgtttgc catcgacatt cttgatcctt cctggtttgt agaacttctc 180  
catctgtcta tgggtgatca cttccatgat atgattgata tgctcataaa ctgtggatat 240  
gagaaaagga ccacactatt tggatatggg tatgattttc gccaaagcaa caggatagac 300  
aaagcgatgg ctggttttgag agcaaaaactt gagacagctc ataagacctc tggagggaaa 360  
aaagttaatt taatctcaca ttctatgggt ggattgctag tacgctgctt catgtctatg 420  
aatcatgatg tattcactaa gtatgtcaac aaatggattt gcattgcttg tccattccaa 480  
ggtgcccccg gatgcatcaa tgattotcta cttactggat tgcaatttgt ttatggtttt 540  
gagagcttct ttttcgtatc tagatgggca atgcaccaat tgcttgtcga atgcccata 600  
atctatgaaa tgttaccaaa tccagaattc aagtggaggg aaaaaccaat tattcaggtt 660  
tggcgtaaga accctgaaaa ggatggaact gtggagcttg ttcaatatga agcaactgat 720  
tgtgtgtcct tgttcgaaga agctttaagg aataatgagc tcacgtataa cggaaagaaa 780  
gtagcactac cattcaatat gtcagtcttc aaatgggcca ccaagactcg ccaaattccta 840  
gacaatgctg aattaccaga tactgtgagc ttttacaata tatacgggac atcttatgaa 900  
actccatacg atgtatgcta tggtcagaa agctctccga ttggagattt gtcagaagtg 960  
tgtcacacag tgccggcata cacttatgtg gatggagatt gcacggttcc catagaatcg 1020  
gcacgggctg atgggttctc tgcgaaagaa agagttggcg tcaaggcgga ccaccgtggc 1080  
ctgctgtccg atgagaacgt attcaagctt ctcaagaaat ggctcgggtg gagcgagaag 1140  
aagtcagagt ggcgttgctg gtctaaatcc tactccaaag tgacctaat gggttgcttg 1200  
tagttcttca ggaagactgt tattttggcc tttcctcctg aagagaagat gaaacaaaat 1260  
tctggtgatt gtattgtatg tctgcacgat gtaaattctt gcaagctgca cggaacaagg 1320  
gattagtgcc cttgtacgat gtatcattgg caggcatttn tttttgaacc tangggcata 1380  
ttnttttgnc cttccactct ggacntagta aagaatatnt gaatcgacct tanttnnaan 1440  
nngtctgnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1500  
nnnnnnnnnn nnaaaaaaaa awgkgaagcc gntnntnntt tnaaaagnnt tttnnnaaaa 1560  
aaaaaaaaaa aaaaaaaaaa aaaaaaaa 1587

<210> 26

<211> 300

<212> DNA

<213> Zea mays

<220>

<221> unsure

<222> (1) .. (300)

<223> n=unknown

<400> 26

gacaaaagcga tggctggttt gagagcaaaa cttgagacag ctcataagac ctctggaggg 60  
aaaaaagtta atttaatctc acattctatg ggtggattgc tagtacgctg cttcatgtct 120  
atgaatcatg atgtgagttt tcatgttttc tgtgtttttt ttgcttttgc ataaatatcc 180  
atgtcaattt cccccatttt ctagggtatc actangtatg tcaacaaatg gatttgcatt 240  
gcttgtccat tccaaggtaa cttatgggac atttcaattg tttattanat natggggncc 300

Variable	Mean	SD	Min	Max
Age	34.5	10.2	18	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	1.5	9	16
Income	3500	1500	1000	8000
Health status	0.8	0.4	0	1
Smoking status	0.3	0.5	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	4.5	1.5	1	7
Sleep quality	3.5	1.0	1	5
Work satisfaction	4.0	1.2	1	5
Life satisfaction	4.2	1.1	1	5
Depression score	2.5	1.5	0	5
Anxiety score	2.0	1.2	0	4
Loneliness score	2.8	1.3	0	5
Self-esteem score	3.0	1.0	1	4
Resilience score	3.2	1.1	1	4
Optimism score	3.5	1.2	1	4
Gratitude score	3.8	1.0	1	4
Forgiveness score	3.6	1.1	1	4
Empathy score	3.4	1.0	1	4
Compassion score	3.3	0.9	1	4
Kindness score	3.1	0.8	1	4
Generosity score	3.0	0.7	1	4
Patience score	2.9	0.6	1	4
Humility score	2.8	0.5	1	4
Modesty score	2.7	0.4	1	4
Shyness score	2.6	0.3	1	4
Introversion score	2.5	0.2	1	4
Extroversion score	2.4	0.1	1	4
Sociability score	2.3	0.1	1	4
Communication score	2.2	0.1	1	4
Interpersonal skills score	2.1	0.1	1	4
Teamwork score	2.0	0.1	1	4
Leadership score	1.9	0.1	1	4
Problem-solving score	1.8	0.1	1	4
Decision-making score	1.7	0.1	1	4
Time management score	1.6	0.1	1	4
Organization score	1.5	0.1	1	4
Productivity score	1.4	0.1	1	4
Efficiency score	1.3	0.1	1	4
Quality of work score	1.2	0.1	1	4
Job satisfaction score	1.1	0.1	1	4
Work-life balance score	1.0	0.1	1	4
Overall well-being score	0.9	0.1	1	4

Variable	Mean	SD	Min	Max
Age	34.5	10.2	18	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	1.5	9	16
Income	3500	1500	1000	8000
Health status	0.8	0.4	0	1
Smoking status	0.3	0.5	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	4.5	1.5	1	7
Sleep quality	3.5	1.0	1	5
Work satisfaction	4.0	1.2	1	5
Life satisfaction	4.2	1.1	1	5
Depression score	2.5	1.5	0	5
Anxiety score	2.0	1.2	0	4
Loneliness score	2.8	1.3	0	5
Self-esteem score	3.0	1.0	1	4
Resilience score	3.2	1.1	1	4
Optimism score	3.5	1.2	1	4
Gratitude score	3.8	1.0	1	4
Forgiveness score	3.6	1.1	1	4
Empathy score	3.4	1.0	1	4
Compassion score	3.3	0.9	1	4
Kindness score	3.1	0.8	1	4
Generosity score	3.0	0.7	1	4
Patience score	2.9	0.6	1	4
Humility score	2.8	0.5	1	4
Modesty score	2.7	0.4	1	4
Shyness score	2.6	0.3	1	4
Introversion score	2.5	0.2	1	4
Extroversion score	2.4	0.1	1	4
Sociability score	2.3	0.1	1	4
Communication score	2.2	0.1	1	4
Interpersonal skills score	2.1	0.1	1	4
Teamwork score	2.0	0.1	1	4
Leadership score	1.9	0.1	1	4
Problem-solving score	1.8	0.1	1	4
Decision-making score	1.7	0.1	1	4
Time management score	1.6	0.1	1	4
Organization score	1.5	0.1	1	4
Productivity score	1.4	0.1	1	4
Efficiency score	1.3	0.1	1	4
Quality of work score	1.2	0.1	1	4
Job satisfaction score	1.1	0.1	1	4
Work-life balance score	1.0	0.1	1	4
Overall well-being score	0.9	0.1	1	4

Variable	Mean	SD	Min	Max
Age	34.5	10.2	18	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	1.5	9	16
Income	3500	1500	1000	8000
Health status	0.8	0.4	0	1
Smoking status	0.3	0.5	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	4.5	1.5	1	7
Sleep quality	3.5	1.0	1	5
Work satisfaction	4.0	1.2	1	5
Life satisfaction	4.2	1.1	1	5
Depression score	2.5	1.5	0	5
Anxiety score	2.0	1.2	0	4
Loneliness score	2.8	1.3	0	5
Self-esteem score	3.0	1.0	1	4
Resilience score	3.2	1.1	1	4
Optimism score	3.5	1.2	1	4
Gratitude score	3.8	1.0	1	4
Forgiveness score	3.6	1.1	1	4
Empathy score	3.4	1.0	1	4
Compassion score	3.3	0.9	1	4
Kindness score	3.1	0.8	1	4
Generosity score	3.0	0.7	1	4
Patience score	2.9	0.6	1	4
Humility score	2.8	0.5	1	4
Modesty score	2.7	0.4	1	4
Shyness score	2.6	0.3	1	4
Introversion score	2.5	0.2	1	4
Extroversion score	2.4	0.1	1	4
Sociability score	2.3	0.1	1	4
Communication score	2.2	0.1	1	4
Interpersonal skills score	2.1	0.1	1	4
Teamwork score	2.0	0.1	1	4
Leadership score	1.9	0.1	1	4
Problem-solving score	1.8	0.1	1	4
Decision-making score	1.7	0.1	1	4
Time management score	1.6	0.1	1	4
Organization score	1.5	0.1	1	4
Productivity score	1.4	0.1	1	4
Efficiency score	1.3	0.1	1	4
Quality of work score	1.2	0.1	1	4
Job satisfaction score	1.1	0.1	1	4
Work-life balance score	1.0	0.1	1	4
Overall well-being score	0.9	0.1	1	4

Variable	Mean	SD	Min	Max
Age	34.5	10.2	18	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	1.5	9	16
Income	3500	1500	1000	8000
Health status	0.8	0.4	0	1
Smoking status	0.3	0.5	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	4.5	1.5	1	7
Sleep quality	3.5	1.0	1	5
Work satisfaction	4.0	1.2	1	5
Life satisfaction	4.2	1.1	1	5
Depression score	2.5	1.5	0	5
Anxiety score	2.0	1.2	0	4
Loneliness score	2.8	1.3	0	5
Self-esteem score	3.0	1.0	1	4
Resilience score	3.2	1.1	1	4
Optimism score	3.5	1.2	1	4
Gratitude score	3.8	1.0	1	4
Forgiveness score	3.6	1.1	1	4
Empathy score	3.4	1.0	1	4
Compassion score	3.3	0.9	1	4
Kindness score	3.1	0.8	1	4
Generosity score	3.0	0.7	1	4
Patience score	2.9	0.6	1	4
Humility score	2.8	0.5	1	4
Modesty score	2.7	0.4	1	4
Shyness score	2.6	0.3	1	4
Introversion score	2.5	0.2	1	4
Extroversion score	2.4	0.1	1	4
Sociability score	2.3	0.1	1	4
Communication score	2.2	0.1	1	4
Interpersonal skills score	2.1	0.1	1	4
Teamwork score	2.0	0.1	1	4
Leadership score	1.9	0.1	1	4
Problem-solving score	1.8	0.1	1	4
Decision-making score	1.7	0.1	1	4
Time management score	1.6	0.1	1	4
Organization score	1.5	0.1	1	4
Productivity score	1.4	0.1	1	4
Efficiency score	1.3	0.1	1	4
Quality of work score	1.2	0.1	1	4
Job satisfaction score	1.1	0.1	1	4
Work-life balance score	1.0	0.1	1	4
Overall well-being score	0.9	0.1	1	4

Variable	Mean	SD	Min	Max
Age	38.5	12.5	18	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	2.5	9	16
Income	3500	1500	1000	8000
Health status	0.8	0.4	0	1
Smoking status	0.3	0.5	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	4.5	1.5	1	7
Sleep quality	3.5	1.5	1	7
Mood stability	4.0	1.5	1	7
Life satisfaction	5.5	1.5	1	7
Work-life balance	4.0	1.5	1	7
Family support	5.0	1.5	1	7
Community involvement	3.0	1.5	1	7
Personal growth	4.5	1.5	1	7
Resilience	5.0	1.5	1	7
Emotional regulation	4.0	1.5	1	7
Interpersonal relationships	4.5	1.5	1	7
Self-awareness	5.0	1.5	1	7
Goal setting	4.0	1.5	1	7
Time management	3.5	1.5	1	7
Decision making	4.5	1.5	1	7
Problem solving	4.0	1.5	1	7
Adaptability	4.5	1.5	1	7
Optimism	5.0	1.5	1	7
Gratitude	4.5	1.5	1	7
Forgiveness	4.0	1.5	1	7
Empathy	4.5	1.5	1	7
Compassion	4.0	1.5	1	7
Kindness	4.5	1.5	1	7
Patience	4.0	1.5	1	7
Perseverance	4.5	1.5	1	7
Discipline	4.0	1.5	1	7
Responsibility	4.5	1.5	1	7
Integrity	4.0	1.5	1	7
Honesty	4.5	1.5	1	7
Trustworthiness	4.0	1.5	1	7
Reliability	4.5	1.5	1	7
Consistency	4.0	1.5	1	7
Stability	4.5	1.5	1	7
Endurance	4.0	1.5	1	7
Strength	4.5	1.5	1	7
Power	4.0	1.5	1	7
Influence	4.5	1.5	1	7
Authority	4.0	1.5	1	7
Leadership	4.5	1.5	1	7
Management	4.0	1.5	1	7
Organization	4.5	1.5	1	7
Coordination	4.0	1.5	1	7
Collaboration	4.5	1.5	1	7
Teamwork	4.0	1.5	1	7
Partnership	4.5	1.5	1	7
Relationship	4.0	1.5	1	7
Connection	4.5	1.5	1	7
Network	4.0	1.5	1	7
Community	4.5	1.5	1	7
Society	4.0	1.5	1	7
World	4.5	1.5	1	7
Universe	4.0	1.5	1	7
Cosmos	4.5	1.5	1	7
Nature	4.0	1.5	1	7
Environment	4.5	1.5	1	7
Worldview	4.0	1.5	1	7
Philosophy	4.5	1.5	1	7
Religion	4.0	1.5	1	7
Spirituality	4.5	1.5	1	7
Metaphysics	4.0	1.5	1	7
Ontology	4.5	1.5	1	7
Epistemology	4.0	1.5	1	7
Axiology	4.5	1.5	1	7
Esthetics	4.0	1.5	1	7
Logic	4.5	1.5	1	7
Mathematics	4.0	1.5	1	7
Science	4.5	1.5	1	7
Technology	4.0	1.5	1	7
Innovation	4.5	1.5	1	7
Progress	4.0	1.5	1	7
Development	4.5	1.5	1	7
Growth	4.0	1.5	1	7
Evolution	4.5	1.5	1	7
Change	4.0	1.5	1	7
Transformation	4.5	1.5	1	7
Reformation	4.0	1.5	1	7
Renewal	4.5	1.5	1	

Descriptive Statistics		Descriptive Statistics		Descriptive Statistics	
Variable	Mean	SD	Variable	Mean	SD
Age	35.2	12.5	Gender	0.5	0.5
Education	12.8	1.5	Marital Status	0.3	0.5
Income	15.5	8.2	Occupation	0.2	0.4
Health Status	2.1	1.2	Stress Level	3.5	1.8
Life Satisfaction	4.2	1.5	Resilience	2.8	1.2
Depression	1.8	1.0	Optimism	3.2	1.5
Anxiety	2.5	1.2	Self-Esteem	3.8	1.8
Loneliness	2.2	1.1	Gratitude	3.5	1.6
Life Purpose	3.0	1.4	Meaning	3.3	1.7
Relationships	2.8	1.3	Community	3.1	1.5
Work-Life Balance	2.5	1.2	Work Satisfaction	3.0	1.4
Financial Stability	2.3	1.1	Healthcare Access	2.9	1.3
Environmental Quality	2.1	1.0	Social Support	2.7	1.2
Personal Growth	2.0	1.0	Life Goals	2.6	1.1
Physical Health	1.9	0.9	Mental Health	2.5	1.0
Emotional Health	1.8	0.8	Spiritual Health	2.4	0.9
Intellectual Health	1.7	0.7	Overall Well-being	2.3	0.8
Social Health	1.6	0.6			
Environmental Health	1.5	0.5			
Personal Health	1.4	0.4			
Relationship Health	1.3	0.3			
Work Health	1.2	0.2			
Financial Health	1.1	0.1			
Environmental Health	1.0	0.0			
Personal Health	0.9	0.0			
Social Health	0.8	0.0			
Environmental Health	0.7	0.0			
Personal Health	0.6	0.0			
Social Health	0.5	0.0			
Environmental Health	0.4	0.0			
Personal Health	0.3	0.0			
Social Health	0.2	0.0			
Environmental Health	0.1	0.0			
Personal Health	0.0	0.0			

```
<210> 30
<211> 518
<212> DNA
<213> Mus musculus
```

```
<210> 31
<211> 299
<212> DNA
<213> Mus musculus
```

```
<210> 32
<211> 1895
<212> DNA
```



Table 1. Demographic characteristics of the study population	
Age (years)	Mean (SD)
18-24	20.5 (2.5)
25-34	29.5 (4.5)
35-44	39.5 (5.5)
45-54	49.5 (6.5)
55-64	59.5 (7.5)
65-74	69.5 (8.5)
75-84	79.5 (9.5)
85-94	89.5 (10.5)
95-104	99.5 (11.5)
105-114	109.5 (12.5)
115-124	119.5 (13.5)
125-134	129.5 (14.5)
135-144	139.5 (15.5)
145-154	149.5 (16.5)
155-164	159.5 (17.5)
165-174	169.5 (18.5)
175-184	179.5 (19.5)
185-194	189.5 (20.5)
195-204	199.5 (21.5)
205-214	209.5 (22.5)
215-224	219.5 (23.5)
225-234	229.5 (24.5)
235-244	239.5 (25.5)
245-254	249.5 (26.5)
255-264	259.5 (27.5)
265-274	269.5 (28.5)
275-284	279.5 (29.5)
285-294	289.5 (30.5)
295-304	299.5 (31.5)
305-314	309.5 (32.5)
315-324	319.5 (33.5)
325-334	329.5 (34.5)
335-344	339.5 (35.5)
345-354	349.5 (36.5)
355-364	359.5 (37.5)
365-374	369.5 (38.5)
375-384	379.5 (39.5)
385-394	389.5 (40.5)
395-404	399.5 (41.5)
405-414	409.5 (42.5)
415-424	419.5 (43.5)
425-434	429.5 (44.5)
435-444	439.5 (45.5)
445-454	449.5 (46.5)
455-464	459.5 (47.5)
465-474	469.5 (48.5)
475-484	479.5 (49.5)
485-494	489.5 (50.5)
495-504	499.5 (51.5)
505-514	509.5 (52.5)
515-524	519.5 (53.5)
525-534	529.5 (54.5)
535-544	539.5 (55.5)
545-554	549.5 (56.5)
555-564	559.5 (57.5)
565-574	569.5 (58.5)
575-584	579.5 (59.5)
585-594	589.5 (60.5)
595-604	599.5 (61.5)
605-614	609.5 (62.5)
615-624	619.5 (63.5)
625-634	629.5 (64.5)
635-644	639.5 (65.5)
645-654	649.5 (66.5)
655-664	659.5 (67.5)
665-674	669.5 (68.5)
675-684	679.5 (69.5)
685-694	689.5 (70.5)
695-704	699.5 (71.5)
705-714	709.5 (72.5)
715-724	719.5 (73.5)
725-734	729.5 (74.5)
735-744	739.5 (75.5)
745-754	749.5 (76.5)
755-764	759.5 (77.5)
765-774	769.5 (78.5)
775-784	779.5 (79.5)
785-794	789.5 (80.5)
795-804	799.5 (81.5)
805-814	809.5 (82.5)
815-824	819.5 (83.5)
825-834	829.5 (84.5)
835-844	839.5 (85.5)
845-854	849.5 (86.5)
855-864	859.5 (87.5)
865-874	869.5 (88.5)
875-884	879.5 (89.5)
885-894	889.5 (90.5)
895-904	899.5 (91.5)
905-914	909.5 (92.5)
915-924	919.5 (93.5)
925-934	929.5 (94.5)
935-944	939.5 (95.5)
945-954	949.5 (96.5)
955-964	959.5 (97.5)
965-974	969.5 (98.5)
975-984	979.5 (99.5)
985-994	989.5 (100.5)
995-1004	999.5 (101.5)
1005-1014	1009.5 (102.5)
1015-1024	1019.5 (103.5)
1025-1034	1029.5 (104.5)
1035-1044	1039.5 (105.5)
1045-1054	1049.5 (106.5)
1055-1064	1059.5 (107.5)
1065-1074	1069.5 (108.5)
1075-1084	1079.5 (109.5)
1085-1094	1089.5 (110.5)
1095-1104	1099.5 (111.5)
1105-1114	1109.5 (112.5)
1115-1124	1119.5 (113.5)
1125-1134	

<223> Description of Artificial Sequence: Inferred cDNA  
sequence

<221> unsure

<223> n=unknown

gtctggtgtg	atggggacag	ggagggactt	ccccctacc	agcactggtg	tgggtgagg	60
tgggtgctga	gtctcagagc	ttggcatgga	gaccagacag	ggctgggtct	gcaagcctga	120
ggctgccgcc	ctgagctcgg	gctgggacgt	gcccagaggt	gttgggagga	tctggggtga	180
gtaccctgtg	gccaggacta	aaggggctnc	accctcctgt	ccatccctcg	cagatcttga	240
gcaatgcccg	gttattttctg	gagaacctca	tcaagtatgg	catcctggtg	gaccccatcc	300
aggtggtttc	tctgttcctg	aaggatccct	atagctggcc	cgcccatgc	ctggttattg	360
cggccaatgt	ctttgctgtg	gctgcattcc	aggttgagaa	gcgcctggcg	gtgggtgcc	420
tgacggagca	ggcgggactg	ctgctgcacg	tggccaacct	ggccaccatt	ctgtgtttcc	480
cagcggctgt	ggtcttactg	gttgagtcta	tactccagt	gggtccctg	ctggcgctga	540
tggcgcacac	catcctcttc	ctcaagctct	tctcctaccg	cgacgtcaac	tcatggtgcc	600
gcagggccag	ggccaaggct	gcctctgcag	ggaagaaggc	cagcagtgt	gctgccccgc	660
acaccgtgag	ctacccgga	aatctgacct	accgcgatct	ctactacttc	ctcttcgccc	720
ccaccttgtg	ctacgagctc	aactttcccc	gctctccccg	catccggaag	cgctttctgc	780
tgcgacggat	ccttgagatg	ctgttcttca	cccagctcca	ggtggggctg	atccagcagt	840
ggatggctcc	caccatccag	aactccatga	agcccttcaa	ggacatggac	tactcacgca	900
tcacgagcgc	cctcctgaag	ctggcggtcc	ccaatcacct	catctggctc	atcttcttct	960
actggctctt	ccactcctgc	ctgaatgccg	tggctgagct	catgcagttt	ggagaccggg	1020
agttctaccg	ggactggtgg	aactccgagt	ctgtcaccta	cttctggcag	aactggaaca	1080
tcctgtgca	caagtgggtc	atcagacact	tctacaagcc	catgcttcga	cggggcagca	1140
gcaagtggat	ggccaggaca	ggggtgttcc	tggcctcggc	cttcttccac	gagtacctgg	1200
tgagcgtccc	tctgcgaatg	ttccgcctct	gggcgttcac	gggcatgatg	gctcagatcc	1260
cactggcctg	gttcgtgggc	cgctttttcc	agggcaacta	tggcaacgca	gctgtgtggc	1320
tgtcgctcat	catcggacag	ccaatagccg	tcctcatgta	cgtccacgac	tactacgtgc	1380
tcaactatga	ggccccagcg	gcagaggcct	gagctgcacc	tgagggcctg	gcttctcact	1440
gccacctcac	acccgctgcc	agagcccacc	tctcctccta	ggcctcgagt	gctggggatg	1500
ggcctggctg	cacagcatcc	tcctctggtc	ccagggaggc	ctctctgccc	ctatggggct	1560
ctgtcctgca	cccctcaggg	atggcgacag	caggccagac	acagtctgat	gccagctggg	1620
agtcttgctg	accctgcccc	gggtccgagg	gtgtcaataa	agtgtgttcc	agtgcacctt	1680
tcagcctgcc	aggggcctgg	ggcctgggtg	ggggtatggc	cacacccaca	agggcgagtg	1740
ccagagctgt	gtggacagct	gtcccaggac	ctgccgggga	gcagcagctc	cactgcagca	1800
gggcgggcat	ggccggtagg	gggagtgcaa	ggccaggcag	acgcccccat	tccccacact	1860
cccctacct	gaaaagctca	gctcaggcgt	cctct			1895

<211> 1766

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Inferred cDNA  
sequence

<400> 33

cacgactggg ccgcgacgtg gtgcgggccg aagccatggg cgaccgcgga ggcgcgggaa 60  
gctctcggcg tcggaggacc ggctcgcggg tttccatcca gggtagtagt gggcccatgg 120  
tagacgaaga ggaggtgcga gacgccgtg tgggccccga cttgggcgcc gggggtgacg 180  
ctccggctcc ggctccggtt ccggctccag cccacacccg ggacaaagac cggcagacca 240  
gcgtgggcga cggccactgg gagctgaggt gccatcgtct gcaagactct ttgttcagct 300  
cagacagcgg tttcagcaat taccgtggtg tcctgaattg gtgctggtg atgctgatcc 360  
tgagtaatgc aaggttatct ttagagaatc ttatcaagta tggcatcctg gtggatccca 420  
tccagggtgg gtctctgttt ctgaaggacc cctacagctg gcctgccccg tgcttgatca 480  
ttgcatccaa tatctttatt gtggctacat ttcagattga gaagcgcctg tcagtgggtg 540  
ccctgacaga gcagatgggg ctgctgctac atgtgggtta cctggccaca attatctgct 600  
tcccagcagc tgtggcctta ctggttgagt ctatcactcc agtgggttcc ctgtttgtct 660  
tggcatcata ctccatcatc ttcctcaagc ttttctccta ccgggatgtc aatctgtggt 720  
gccgccagcg aagggtcaag gccaaagctg tgtctgcagg gaagaaggtc agtggggctg 780  
ctgccagaa cactgtaagc tatccggaca acctgacctc ccgagatctc tattacttca 840  
tctttgtctc tactttgtgt tatgaactca actttcctcg atccccccga atacgaaagc 900  
gctttctgct acggcggggt cttgagatgc tctttttcac ccagcttcaa gtggggctga 960  
tccagcagtg gatggtccct actatccaga actccatgaa gcccttcaag gacatggact 1020  
attcacgaat cattgagcgt ctcttaaagc tggcgggtccc caaccatctg atatggctca 1080  
tcttcttcta ttggcttttc cactcatgtc tcaatgctgt ggagagctc ctgcagtttg 1140  
gagaccgca gttctacagg gactggtgga atgctgagtc tgtcacctac ttttggcaga 1200  
actggaatat ccccgtcac aagtgggtgca tcagacactt ctacaagcct atgctcagac 1260  
tgggcagcaa caaatggatg gccaggactg ggggtctttt ggcgtcagcc ttcttccatg 1320  
agtacctagt gagcattccc ctgaggatgt tccgcctctg ggcattcaca gccatgatgg 1380  
ctcaggtccc actggcctgg attgtgaacc gcttcttcca agggaactat ggcaatgcag 1440  
ctgtgtgggt gacactcatc attgggcaac cgggtggctgt gctcatgtat gtccacgact 1500  
actacgtgct caactatgat gcccagtggt gggcctgagc tactgccaaa ggccagccct 1560  
ccctaacctg ggctggagt tctggagggc ttctggtgct cctgcacact cctcctagtc 1620  
tgggaggcct ctctgcccc atggggccta ctctgctct tggggatggc acctgagtcc 1680  
agctggtatg agccagtgt gggagtctgt gctgaccagg ggctgaggat atcaataaag 1740  
agctatctaa aaaaaaaaa aaaaaa 1766

<210> 34

<211> 409

<212> PRT

<213> Homo sapiens

<400> 34

Arg Arg Ser Leu Leu Asp Glu Leu Leu Glu Val Asp His Ile Arg Thr

1

5

10

15

















[illegible]

ctctcgtgaa	tcttttttcc	ttctttcttc	ttctttctct	cagagaaaac	tttgcttctc	60
tttctataag	gaaccagaca	cgaatcccat	tcccaccgat	ttcttagctt	cttccttcaa	120
tccgctcttt	cctctcccat	tagattctgt	ttctcttttc	aattttcttct	gcatgcttct	180
cgattctctc	tgacgcctct	tttctcccga	cgctgtttcg	tcaaacgctt	ttcgaaatgg	240
cgatttttga	ttctgctggc	gttactacgg	tgacggagaa	cgggtggcga	gagttcgtcg	300
atcttgatag	gcttcgtcga	cggaaatcga	gacggatttc	ttctaacgga	cttcttctct	360
ctggttccga	taataattct	ccttcggatg	atgttgagac	tcccgcgcac	gttagggatc	420
ggattgattc	cgttgttaac	gatgacgctc	agggaaacagc	caatttggcc	ggagataata	480
acggtgggtg	cgataataac	ggtggtggaa	gaggcggcgg	agaaggaaga	ggaaacgccg	540
atgctacgtt	tacgtatcga	ccgtcggttc	cagctcatcg	gagggcgaga	gagagtccac	600
ttagctccga	cgcaatcttc	aaacagagcc	atgccggatt	attcaacctc	tgtgtagtag	660
ttcttattgc	tgtaaacagt	agactcatca	tcgaaaatct	tatgaagtat	ggttggttga	720
tcagaacgga	tttctggttt	agttcaagat	cgctgcgaga	ttggcgcgtt	ttcatgtgtt	780
gtatatccct	ttcgatcttt	cctttggctg	cctttacggt	tgagaaattg	gtacttcaga	840
aatacatatc	agaacctgtt	gtcatctttc	ttcatattat	tatcaccatg	acagaggttt	900
tgtatccagt	ttacgtcacc	ctaagggtgtg	attctgcttt	tttatcaggt	gtcactttga	960
tgctcctcac	ttgcattgtg	tggctaaaagt	tggttttctta	tgtctatact	agctatgaca	1020
taagatccct	agccaatgca	gctgataagg	ccaatcctga	agtctcctac	tacgttagct	1080
tgaagagctt	ggcatatttc	atggtcgcctc	ccacattgtg	ttatcagcca	agttatccac	1140
gttctgcatg	tatacggaaag	ggttggggtg	ctcgtcgaatt	tgcaaaaactg	gtcatattca	1200
ccggattcat	gggatttata	atagaacaat	atataaatcc	tattgtcagg	aactcaaagc	1260
atcctttgaa	aggcgatctt	ctatatgcta	ttgaaagagt	gttgaagctt	tcagttccaa	1320
atztatatgt	gtggctctcg	atgttctact	gcttcttcca	cctttgggtta	aacatattgg	1380
cagagcttct	ctgcttcggg	gacgtgaat	tctacaaaga	ttggtggaat	gcaaaaagtg	1440
tgggagatta	ctggagaatg	tggaaatatgc	ctgttcataa	atggatgggt	cgacatatat	1500
acttcccgtg	cttgccgcagc	aagataccaa	agacactcgc	cattatcatt	gctttcctag	1560
tctctgcagt	ctttcatgag	ctatgcatcg	cagttccttg	togtctcttc	aagctatggg	1620
cttttcttgg	gattatgttt	cagggtgcctt	tgggtcttcat	cacaaactat	ctacaggaaa	1680
ggtttggctc	aacggtgggg	aacatgatct	tctggttcat	cttctgcatt	ttcggacaac	1740
cgatgtgtgt	gcttctttat	taccacgacc	tgatgaaccg	aaaaggatcg	atgtcatgaa	1800
acaactgttc	aaaaaatgac	tttcttcaaa	catctatggc	ctcgttggat	ctccgttgat	1860
gttggtgggtg	ttctgatgct	aaaacgacaa	atagtgttat	aaccattgaa	gaagaaaaga	1920
caattagagt	tgttgtatcg	ca				1942

<211> 520











atgagttatg tgttgctggt ccttgccaca tactcaagtt ctgggctttt tttaggaatc 180  
 atgcttcaga ttccctcat catattgaca tcatacctca aaaataaaatt cagtgcacaca 240  
 atggttggca ata 253

<210> 48  
 <211> 254  
 <212> DNA  
 <213> Zea mays

<400> 48  
 tgaagtatgg cttattaata agatctggct tttggtttaa tgctacatca ttgcgagact 60  
 ggccactgct aatgtgttgc cttagtctac ccatatttcc cttggtgca ttgcagtcg 120  
 aaaagttggc attcaacaat ctcattagtg atcctgctac tacctgtttt cacatccttt 180  
 ttacaacatt tgaaattgta tatccagtgc tcgtgattct taagtgtgat tctgcagttt 240  
 tatcaggctt tgtg 254

<210> 49  
 <211> 262  
 <212> DNA  
 <213> Zea mays

<400> 49  
 gaagtatggc ttattaataa gatctggctt ttggtttaat gctacatcat tgcgagactg 60  
 gccactgcta atgtgttggc ttagtctacc catatttccc cttggtgcat ttgcagtcga 120  
 aaagttggca ttcaacaatc tcattagtga tcctgctact acctgttttc acatcctttt 180  
 tacaacattt gaaattgtat atccagtgc cgtgattctt aagtgtgatt ctgcagtttt 240  
 acaggctttg tgttgatgtt ta 262

<210> 50  
 <211> 325  
 <212> DNA  
 <213> Zea mays

<220>  
 <221> unsure  
 <222> (1) .. (325)  
 <223> n=unknown

<400> 50  
 taatcnaacc tcgntncngg ttcagctgta tnccatgaga tatgtaatgc ggtgccgtgc 60  
 cacatantca natctnggca tnnnnggat catngttcag ataccgntgg nattcttgac 120  
 aagatatctc catgctacgt tcaagcatgt aatgggtgggc aacatgatan tttggntctn 180  
 cagtatagtc ggacagccga tgnnnnnna tctatactac catgacgtca tgaacaggca 240  
 ggcccaggca agtagatagt ncggcagaga catgtacttc aacatcganc atcagnagca 300  
 nacngagcga gcggcangaa ncage 325



Socioeconomic characteristics	
Age (years)	40.0
Gender	Male
Marital status	Married
Education (years)	12.0
Occupation	Unemployed
Income (USD/month)	1000
Health insurance	Yes
Smoking status	Non-smoker
Alcohol consumption	None
Exercise frequency	None
Stress level	High
Family size	3
Urban/rural	Urban
Season	Spring
Time of day	Daytime
Weather	Sunny
Location	Indoor
Time of year	Summer
Month	June
Day	15
Hour	14:00
Minute	30
Second	00
Millisecond	000
Microsecond	0000
Nanosecond	000000
Picosecond	00000000
Femtosecond	0000000000
Attosecond	000000000000
Zeptosecond	00000000000000
Yoctosecond	0000000000000000
Planck time	5.39e-44
Age (years)	40.0
Gender	Male
Marital status	Married
Education (years)	12.0
Occupation	Unemployed
Income (USD/month)	1000
Health insurance	Yes
Smoking status	Non-smoker
Alcohol consumption	None
Exercise frequency	None
Stress level	High
Family size	3
Urban/rural	Urban
Season	Spring
Time of day	Daytime
Weather	Sunny
Location	Indoor
Time of year	Summer
Month	June
Day	15
Hour	14:00
Minute	30
Second	00
Millisecond	000
Microsecond	0000
Nanosecond	000000
Picosecond	00000000
Femtosecond	0000000000
Attosecond	000000000000
Zeptosecond	00000000000000
Yoctosecond	0000000000000000
Planck time	5.39e-44

<400> 53  
ggatccccctg caggtcattc attgacggca ttaacattgg 40

<210> 54  
<211> 44  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic  
oligonucleotide primer

<400> 54  
ggatccgcg cgacacaatg ggagcgaatt cgaaatcagt aacg 44

<210> 55  
<211> 40  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic  
oligonucleotide primer

<400> 55  
ggatcccctg caggttaata cccactttta tcaagctccc 40

<210> 56  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic  
oligonucleotide primer

<400> 56  
ggatccgcg cgacacaatg tctctattac tggaagagat c 41

<210> 57  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic

Variable	Mean	SD	Min	Max
Age	34.5	10.2	18	65
Gender	0.5	0.5	0	1
Marital status	0.6	0.5	0	1
Education	12.5	1.5	9	16
Income	15.2	5.8	5	35
Occupation	1.2	0.8	0	3
Health status	2.1	0.9	1	4
Stress level	3.5	1.2	1	5
Life satisfaction	4.2	0.8	3	5
Resilience	2.8	1.1	1	5
Optimism	3.1	1.0	1	5
Self-efficacy	3.8	1.3	1	5
Emotional stability	2.5	0.9	1	4
Prosocial behavior	3.3	1.1	1	5
Empathy	3.6	1.2	1	5
Altruism	3.4	1.1	1	5
Cooperativeness	3.2	1.0	1	5
Helpfulness	3.5	1.1	1	5
Kindness	3.7	1.2	1	5
Generosity	3.9	1.3	1	5
Compassion	4.0	1.4	1	5
Forgiveness	4.1	1.5	1	5
Patience	4.2	1.6	1	5
Understanding	4.3	1.7	1	5
Acceptance	4.4	1.8	1	5
Openness	4.5	1.9	1	5
Conscientiousness	4.6	2.0	1	5
Agreeableness	4.7	2.1	1	5
Neuroticism	4.8	2.2	1	5
Extraversion	4.9	2.3	1	5
Introversion	5.0	2.4	1	5
Shyness	5.1	2.5	1	5
Loneliness	5.2	2.6	1	5
Isolation	5.3	2.7	1	5
Depression	5.4	2.8	1	5
Anxiety	5.5	2.9	1	5
Stress	5.6	3.0	1	5
Worry	5.7	3.1	1	5
Overthinking	5.8	3.2	1	5
Self-doubt	5.9	3.3	1	5
Low self-esteem	6.0	3.4	1	5
Loneliness	6.1	3.5	1	5
Isolation	6.2	3.6	1	5
Depression	6.3	3.7	1	5
Anxiety	6.4	3.8	1	5
Stress	6.5	3.9	1	5
Worry	6.6	4.0	1	5
Overthinking	6.7	4.1	1	5
Self-doubt	6.8	4.2	1	5
Low self-esteem	6.9	4.3	1	5
Loneliness	7.0	4.4	1	5
Isolation	7.1	4.5	1	5
Depression	7.2	4.6	1	5
Anxiety	7.3	4.7	1	5
Stress	7.4	4.8	1	5
Worry	7.5	4.9	1	5
Overthinking	7.6	5.0	1	5
Self-doubt	7.7	5.1	1	5
Low self-esteem	7.8	5.2	1	5
Loneliness	7.9	5.3	1	5
Isolation	8.0	5.4	1	5
Depression	8.1	5.5	1	5
Anxiety	8.2	5.6	1	5
Stress	8.3	5.7	1	5
Worry	8.4	5.8	1	5
Overthinking	8.5	5.9	1	5
Self-doubt	8.6	6.0	1	5
Low self-esteem	8.7	6.1	1	5
Loneliness	8.8	6.2	1	5
Isolation	8.9	6.3	1	5
Depression	9.0	6.4	1	5
Anxiety	9.1	6.5	1	5
Stress	9.2	6.6	1	5
Worry	9.3	6.7	1	5
Overthinking	9.4	6.8	1	5
Self-doubt	9.5	6.9	1	5
Low self-esteem	9.6	7.0	1	5
Loneliness	9.7	7.1	1	5
Isolation	9.8	7.2	1	5
Depression	9.9	7.3	1	5
Anxiety	10.0	7.4	1	5
Stress	10.1	7.5	1	5
Worry	10.2	7.6	1	5
Overthinking	10.3	7.7	1	5
Self-doubt	10.4	7.8	1	5

ggatcccctg caggttatgc atcaacagag aactttacag c

41

<211> 41

<213> Artificial Sequence

<223> Description of Artificial Sequence: Synthetic oligonucleotide primer

ggatccgcgg cgcacaatg ggctggattc cgtgtccgtg c

41

<211> 38

<213> Artificial Sequence

<223> Description of Artificial Sequence: Synthetic  
oligonucleotide primer

ggatcccctg caggttaacc agaatcaact actttgtg

38

<211> 39

### <213> Artificial Sequence

<223> Description of Artificial Sequence:  
Oligonucleotide primer

tcgacctgca ggaagcttag aaatggcgat tttggattc

39

<211> 36

<212> DNA

Variable	Mean	SD	Min	Max
Age	34.5	10.2	21	55
Gender	1.2	0.4	1	2
Marital status	1.5	0.5	1	3
Education	12.5	1.5	9	16
Income	1.8	0.8	1	3
Occupation	1.5	0.5	1	3
Health status	1.5	0.5	1	3
Stress level	2.5	0.8	1	4
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life satisfaction	3.5	0.8	1	5
Resilience	2.5	0.8	1	4
Optimism	3.5	0.8	1	5
Gratitude	3.5	0.8	1	5
Self-esteem	3.5	0.8	1	5
Empathy	3.5	0.8	1	5
Prosocial behavior	3.5	0.8	1	5
Life				

<223> Description of Artificial Sequence:  
Oligonucleotide primer

<400> 61

ggatccgcg cgcgtcatga catcgatcct tttcgg

36

<210> 62

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Annealed oligonucleotide adapter

<400> 62

cgcgatttaa atggcgcgcc ctgcaggcgg ccgcctgcag ggcgcgccat ttaa

56

<210> 63

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ligating oligonucleotide

<400> 63

tcgaggatcc gcggccgcaa gcttcctgca gg

32

<210> 64

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ligating oligonucleotide

<400> 64

tcgacctgca ggaagcttgc ggccgcggat cc

32







1990-1991		1991-1992		1992-1993		1993-1994		1994-1995		1995-1996		1996-1997		1997-1998		1998-1999		1999-2000		2000-2001		2001-2002		2002-2003		2003-2004		2004-2005		2005-2006		2006-2007		2007-2008		2008-2009		2009-2010		2010-2011		2011-2012		2012-2013		2013-2014		2014-2015		2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023		2023-2024		2024-2025		2025-2026		2026-2027		2027-2028		2028-2029		2029-2030		2030-2031		2031-2032		2032-2033		2033-2034		2034-2035		2035-2036		2036-2037		2037-2038		2038-2039		2039-2040		2040-2041		2041-2042		2042-2043		2043-2044		2044-2045		2045-2046		2046-2047		2047-2048		2048-2049		2049-2050		2050-2051		2051-2052		2052-2053		2053-2054		2054-2055		2055-2056		2056-2057		2057-2058		2058-2059		2059-2060		2060-2061		2061-2062		2062-2063		2063-2064		2064-2065		2065-2066		2066-2067		2067-2068		2068-2069		2069-2070		2070-2071		2071-2072		2072-2073		2073-2074		2074-2075		2075-2076		2076-2077		2077-2078		2078-2079		2079-2080		2080-2081		2081-2082		2082-2083		2083-2084		2084-2085		2085-2086		2086-2087		2087-2088		2088-2089		2089-2090		2090-2091		2091-2092		2092-2093		2093-2094		2094-2095		2095-2096		2096-2097		2097-2098		2098-2099		2099-2100		2100-2101		2101-2102		2102-2103		2103-2104		2104-2105		2105-2106		2106-2107		2107-2108		2108-2109		2109-2110		2110-2111		2111-2112		2112-2113		2113-2114		2114-2115		2115-2116		2116-2117		2117-2118		2118-2119		2119-2120		2120-2121		2121-2122		2122-2123		2123-2124		2124-2125		2125-2126		2126-2127		2127-2128		2128-2129		2129-2130		2130-2131		2131-2132		2132-2133		2133-2134		2134-2135		2135-2136		2136-2137		2137-2138		2138-2139		2139-2140		2140-2141		2141-2142		2142-2143		2143-2144		2144-2145		2145-2146		2146-2147		2147-2148		2148-2149		2149-2150		2150-2151		2151-2152		2152-2153		2153-2154		2154-2155		2155-2156		2156-2157		2157-2158		2158-2159		2159-2160		2160-2161		2161-2162		2162-2163		2163-2164		2164-2165		2165-2166		2166-2167		2167-2168		2168-2169		2169-2170		2170-2171		2171-2172		2172-2173		2173-2174		2174-2175		2175-2176		2176-2177		2177-2178		2178-2179		2179-2180		2180-2181		2181-2182		2182-2183		2183-2184		2184-2185		2185-2186		2186-2187		2187-2188		2188-2189		2189-2190		2190-2191		2191-2192		2192-2193		2193-2194		2194-2195		2195-2196		2196-2197		2197-2198		2198-2199		2199-2200		2200-2201		2201-2202		2202-2203		2203-2204		2204-2205		2205-2206		2206-2207		2207-2208		2208-2209		2209-2210		2210-2211		2211-2212		2212-2213		2213-2214		2214-2215		2215-2216		2216-2217	
-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--

1 1 1 1

<213> Arabidopsis thaliana

Val Leu Ile Ala Asn Leu Ala His Ile Gly Tyr Glu Glu Lys Asn Met  
195 200 205















